ERGONOMICS

HUMAN FACTORS IN WORK, MACHINE CONTROL AND EQUIPMENT DESIGN

A Taylor and Francis International Journal

The Official Publication of the Ergonomics Research Society

Volume 3 **Number 3**

July 1960

ALERE FLAMMAM.

Printed and Published by

TAYLOR & FRANCIS LTD.

RED LION COURT, FLEET STREET, LONDON, E.C.4

66

ERGONOMICS

Human Factors in Work, Machine Control and Equipment Design

Price 25s. 0d. per part

Subscription price per volume £4 15s. 0d. post free, payable in advance

Agents for U.S.A. and Canada

Academic Press Inc., 111 Fifth Avenue, New York 3, N.Y., U.S.A.

$3.50 per part. $13.30 per volume.

THE MEASUREMENT OF SENSORY-MOTOR PERFORMANCE: SURVEY AND REAPPRAISAL OF TWELVE YEARS' PROGRESS

By A. T. Welford

The Psychological Laboratory, Cambridge

In recent years the importance of perceptual and central organizing activities in sensory-motor performance has been increasingly recognised and progress has been made towards a genuinely quantitative treatment. This paper sketches the historical development of the work in this area and attempts a reappraisal under five main heads :

(a) There appears to be in the central mechanisms a ' single channel ' which deals with signals or groups of signals one at a time so that signals coming in rapid succession may have to ' queue ' before they are dealt with.

(b) Choice reaction times are discussed in relation to the theory that the subject gains information, in the information-theory sense of the term, at a constant rate. Conceptual models of the subject's detailed behaviour when making choices are also considered.

(c) Information theory models relating to the speed and accuracy of movement are outlined and discussed.

(d) Several formulae attempting to relate time taken to discriminate quantities of different magnitudes and the fineness of the difference between them are examined.

(e) A number of wider implications of the work surveyed are outlined. Perhaps the most important of these are new approaches to ' mental ' and monitoring tasks which have so far not been amenable to the normal methods of work study.

It is concluded that there is a need for joint psychological and physiological research which would be able to go beyond descriptive mathematical formulae to the study of detailed micro-behaviour and neuro-muscular mechanisms.

§ 1. Historical Introduction

A student beginning psychology today must find it very difficult to realize fully the revolution that has taken place in the subject during the last decade. Thought has become clearer and more confident in many areas, and substantial progress has been made towards a psychology which is truly quantitative, and thus towards the removal of what had previously been one of the subject's most serious drawbacks. A physiologist can speak of human performance in terms of body temperatures and pulse rates and relate these to foot-pounds, calories and heat-loads. Hitherto a psychologist has been virtually unable to measure performance in ways which transcend the particular cases concerned. His ability to do so now is still fragmentary, but holds very considerable promise for the future. This paper will sketch the main outlines of this development.

The years 1947 to 1949 produced a series of remarkable publications on sensory-motor performance. First mention should go to two posthumous papers by Craik (1947, 1948) on the Theory of the Human Operator of Control Mechanisms, and an important theoretical paper by Hick (1948). Coupled with Craik's papers was a series of experimental studies by Vince (1948 a, b; 1949) illustrating the theoretical points. On the applied side were books by Fitts (1947) and by Chapanis *et al.* (1949) bringing together several series of researches showing ways in which the design of instruments and other displays

could profoundly influence either the speed, accuracy or both of an operator's performance. They thus emphasized the importance of perceptual aspects of machine and work-design and the fact that their effects are measurable.

The developments were not, of course, without their antecedents. When one reads some of the papers on human performance published in psychological journals towards the end of the nineteenth century one cannot help feeling that all this might have come much earlier but for the controversies of narrow-minded 'schools' of psychology during the first forty years of the present century.

Although the idea of relating measures of performance to physical units was overshadowed during these years it was not extinguished. The opportunity came for it to expand again during the second world war when experimental psychology was challenged to understand human performance and behaviour under a range of conditions that had not been systematically studied before.

The effects of the war-time change of interest might have been far less profound had there not appeared two books, neither specifically concerned with psychology, but destined to have a profound effect upon it. The first was Wiener's *Cybernetics* (1948) which set out ideas, clothed with the appropriate mathematics, for considering man as a self-regulating mechanism, and outlined the basic concepts of 'information theory'. The second was *The Mathematical Theory of Communication* (Shannon and Weaver 1949), which gave an expanded treatment of information theory including a set of theorems. It is often doubted whether information theory has given significant answers to psychological questions, but it has certainly provided a vigorous impetus to conceive of mental processes in quantitative terms.

A sign of the developing thought of the war years was the emphasis by Craik (1943) of the inadequacy of conceiving the brain to act either as a vast telephone exchange of reflex arcs or as a vaguely defined field of interacting forces. Rather it must be thought of as a computer, receiving inputs from many sources and combining them together to produce an output which is unique to the particular occasion although nevertheless lawful.

Electronic computers were 'in the air' during and just after the war, and the idea caught on, albeit often in rather crude interpretations. Its essential significance has been to point out that, among the wide variety of human skilled performances, it is worth looking for a relatively few fundamental mechanisms which would enable us to by-pass or substantially shorten a great deal of *ad hoc* study of particular human operations, whether these studies are made in industry, in the armed services or in the experimental laboratory.

This line of approach led to a series of questions, of which two are of fundamental importance. Firstly, what is the capacity of the computer in terms of the amount of data it can handle at once, and the amount it can deal with in any given period of time? Secondly, the computer must obviously perform several functions ; what are these and how do they interact ?

We shall consider some of the answers that have appeared to these questions during the last twelve years, what picture they give us of human performance, and what practical implications this has today.

§ 2. Capacity of the Human 'Information Channel'

The modern concept of psychological capacity in relation to human performance has developed hand in hand with a revival of interest in reaction times. These were extensively studied in the nineteenth century, but were neglected between the wars to such an extent that only one major textbook of experimental psychology contained any comprehensive treatment of them (Woodworth 1938).

Reaction time includes the time required for the sense organ receiving the stimulus to act and for the responding muscles to be set in motion, but most of it is taken up by processes in the central brain mechanisms. An excellent summary of the classical studies in this field is given by Woodworth (1938) and by Woodworth and Schlosberg (1954). Recent work has been mainly concentrated on five lines which we shall examine in turn.

2.1. *The Single-channel Hypothesis*

Researches by Vince (1948 a, 1950) and others since her (see Welford 1952, 1959; also Broadbent 1958) have indicated that when one signal occurs very shortly after another the time taken to respond to the second signal may be longer than normal. Of various explanations put forward the most attractive and best supported is the view that some part of the central mechanisms can deal with only one signal at a time. If, therefore, a signal appears during the reaction time to a previous signal, the second signal has, as it were, to queue up until the central mechanisms are free. The only important exception to this principle yet discovered is that when two signals occur almost simultaneously—say up to about 0·05 sec apart—they may be dealt with together. In this case, however, a coordinated response appears to be made to both signals as a single 'unit'.

The central mechanisms may be occupied in dealing not only with signals from outside, but also in monitoring actions. Signals may, therefore, have to queue if they arrive during or shortly after the movement in response to a previous signal as they may clash with kinaesthetic or other sensory 'feedback' from the movement. Monitoring of responses may in some cases be eliminated (Davis 1956; Marill 1957) by intensive practice. The reason is presumably that practice enables simple repetitive actions to be made with a degree of precision such that their outcome is not in doubt, and thus no monitoring is necessary. The effects of such monitoring are not only illustrated in experiments on reaction times. Broadbent (1952), for example, using relatively untrained subjects found that attempting to listen to one message while replying to another impaired the accuracy of both listening and replying.

We can thus say that the central mechanisms behave as if they contained at some point a *single channel* dealing with only one signal, or one set of coordinated signals, at a time. The theoretical and practical implications are illustrated by the fact that, in two-handed work, the hands cannot be treated as independent units. They may perform different actions, but these must be coordinated if they are to be strictly simultaneous. To put it another way, there should be on the so-called SIMO charts used by work-study engineers to plot the actions in two-handed work, not only columns for the two hands but a third headed 'Brain : single channel' in which there should be an entry for every independent action of *either* hand.

2.2. *Stages of Central Activity*

Further evidence makes it clear that the central mechanisms do not act as a single whole, but as a chain with at least three links. Firstly, it is known that a response may be made to a signal even though it comes and goes again during the reaction time to a previous signal. This argues that a signal can be received and the data from it stored while a previous signal is being dealt with. Secondly, there is clear evidence that the reception of one signal and the *selection* of a response to it can in certain circumstances overlap with the *execution* of the response to a previous signal. The evidence, taken as a whole, suggests a chain of mechanisms something like that shown in Fig. 1. Stimuli from the display impinge upon the sense organs which convert them into patterns of nerve impulses. These are relayed to a *perceptual* mechanism in which integration and identification take place. At the other end of the chain the effector organs, such as hands or feet, are activated by impulses from central mechanisms which determine the coordination and phasing of muscular action. Connecting the two ends is what may be tentatively called a *translation* mechanism concerned with the choice of action in relation to what is perceived. This receives data previously coordinated at the perceptual stage and passes 'orders' to the central effector stage which then carries out chains of coordinated action.

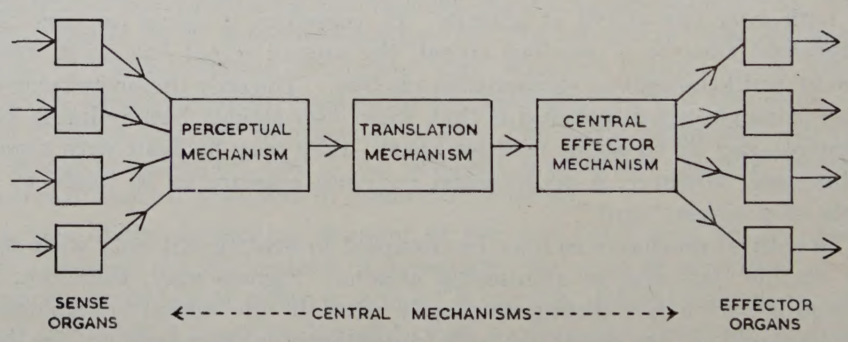

Figure 1. Tentative block diagram of the chain of mechanisms involved in sensory-motor performance.

The manner in which signals pass from one stage to the next is not as yet understood, but the evidence suggests that a feedback loop from the effector side controls the passage of data from the perceptual to the translation stages, letting through data from a new signal once the action in response to the previous signal has begun. The picture is complicated by the fact that delays may be produced by signals which are observed but to which no response is required (Fraisse 1957; Davis 1959). In these cases it is possible that some sub-overt response is nevertheless made or that the inhibition of response is signalled back to the perceptual stage.

Studies of the effects of signals arriving in quick succession leave open the question of *how much* time is taken by each of the mechanisms concerned. The times required by the sense organs to function are relatively short except when working near threshold, and the times added by the peripheral effector organs to those needed for the central coordination of action are also probably short in most cases. The essential question is, therefore, concerned with the times taken by the several central processes.

2.3. *Choice Reaction Times*

Perhaps the most important pioneer break into this problem was made by Hick (1952 a) who proposed, on the basis of his own data and also those of Merkel (1885), that in making choice reactions the subject gains 'information', in the information-theory sense of the term, at a constant rate.

Merkel had presented his subjects with signals ranging in different trials from one to ten alternatives. The signals consisted of the arabic numerals 1–5 and roman numerals I–V, printed round the edge of a disc. The subject waited for each signal with his fingers pressed on ten keys and, when a number was illuminated, released the corresponding key. The arabic numerals corresponded in order to the fingers of the right hand, and the roman to the left. When less than ten choices were required some of the numerals were omitted.

Hick's own experiments used as a display ten pea-lamps arranged in a 'somewhat irregular circle'. The subject reacted by pressing one of ten morse keys on which his fingers rested. Choices of less than ten were again obtained by omitting some of the lights. The frequencies of the various signals for any given degree of choice were carefully balanced and presented in an irregular order so as to ensure as far as possible that the subject should not be able to predict what signal was coming next. Each light appeared 5 sec after the completion of the previous response—an interval too long for the subject to judge accurately when the signal would appear.

Hick found that if the number of possible signals is taken as n and reaction time is plotted against log $(n+1)$, the observed reaction times for different numbers of signals lie on a straight line which also passes through the origin, as shown in Fig. 2. We can thus write

$$\text{Choice reaction time} = K \log (n+1) \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

where K is a constant. If we work in logarithms to the base 2, log $(n+1)=1$ when $n=1$ and K is the simple reaction time—a convenient result.

The obvious question arises, why $(n+1)$ and not n? Hick has pointed out that if the subject is uncertain when a signal will appear he is faced with the task, when it does appear, not only of deciding which it is, but also of deciding that a signal has occurred at all: failure to do so will result in his either reacting when there is no signal present or failing to react when there is one. The additional task of guarding against such errors can be conceived as adding one to the number of possible states of affairs that he has to distinguish—instead of states corresponding to signals 1, 2, 3, . . . n he has to deal with states corresponding to 0, 1, 2, 3, . . . n. If the subject were in no doubt when a signal was coming, as, for example, if he were to determine the point of time himself at which the signal light came on, the $+1$ in eqn. (1) would not be required since there would be no temporal uncertainty to be resolved. We may denote the sum of the possibilities including 'no signal' as N, defining N as the *equivalent total* number of equally probable alternatives from which the subject has to choose, and may then rewrite eqn. (1) as

$$\text{Choice reaction time} = K \log N. \qquad . \quad . \quad . \quad . \quad . \quad (2)$$

This formulation we may call *Hick's Law*. It should be understood that it is an 'ideal' formula and that time lags in the apparatus or in the making of a response may add a constant to the time.

Some confirmation of Hick's view is obtained from an experiment by Crossman (1956 a) who managed to eliminate temporal uncertainty in a task which, in different trials, presented 1, 2, 4 and 8 choices. The subject brought on the signal himself by raising his finger from a 'home' key from which he had to move to one of a row of buttons numbered 1–8 in order a few inches further from him than the key. The signal lights consisted of small windows which when illuminated showed a number similar to that of the corresponding button. They were, in one condition ('symbolic'), in scattered positions on a board behind the buttons so that the subject had to use the number symbol to translate from signal to response. In a second condition ('non-symbolic') the signal lights were placed directly above the buttons so that no symbolic translation was required.

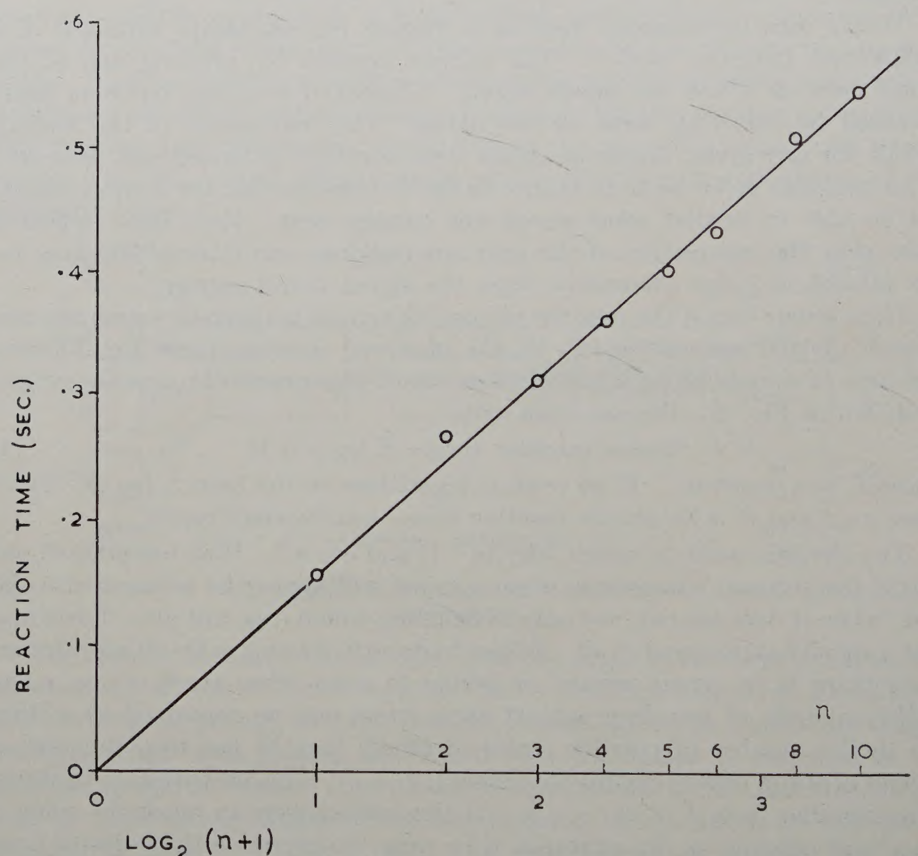

Figure 2. Data from a choice reaction experiment by Hick, plotted in terms of eqn. (1). The total number of reactions represented is over 2400, recorded after extensive practice.

The results are shown in Fig. 3. Since there was no chance of premature or omitted responses, reaction times have been plotted against log n instead of log $(n+1)$. The fit is reasonably close considering that the numbers of reaction times taken were relatively small. The substantial difference of slope between the symbolic and non-symbolic conditions is a matter to which we shall have to return later.

The regression lines intersect approximately at the zero information point ($\log n = 0$), and if eqn. (2) applied straightforwardly this point should also be that for zero time. The time for zero information is, however, not truly a reaction time, but represents the time taken to move from the 'home' key and to press the button, plus any delays of recording in the apparatus. Crossman's results are, therefore, *not* correctly represented by a formulation proposed by Hyman (1953) and also by Bricker (1955) as an alternative to Hick's.

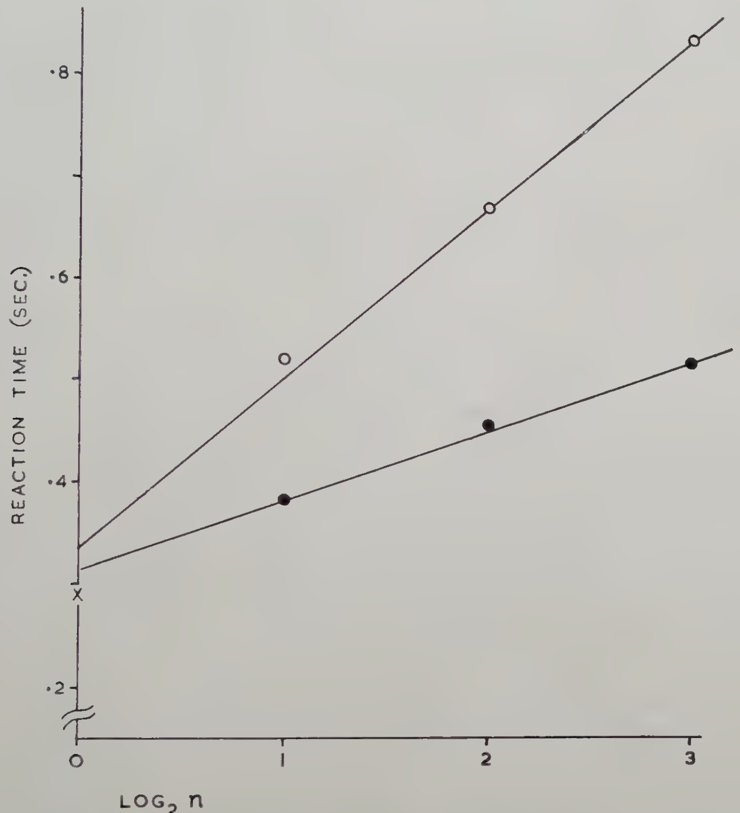

Figure 3. Data from an experiment by Crossman comparing performances with symbolic and non-symbolic displays. Each point is the mean of 160 readings—40 from each of four subjects. Open circles = symbolic displays, filled circles = non-symbolic, × = zero information point.

Hyman's and Bricker's view is that uncertainty about when a signal will come and about which signal has arrived should be treated separately and that we should write instead of eqn. (2)

$$\text{Choice reaction time} = a + b \log n \qquad (3)$$

where a and b are constants, and a is equal to the simple reaction time and caters for temporal uncertainty. Again there may be an additional constant due to time lags in the apparatus or in the execution of the response.

Support for this approach appears, at first sight, to come from results by Klemmer (1957) which indicate that information due to temporal uncertainty in a simple reaction time task can be calculated in terms of the variability

of the warning period given of a signal's appearance, and of the subject's ability to estimate time. Klemmer's calculations suggest a mechanism different from that involved in making choices, and one which gains information at a much higher rate. The variations in reaction time that he reports between different degrees of temporal uncertainty are, however, small and, as Hick has pointed out in a private communication, should be distinguished from the effects of having to guard against premature or omitted reactions. Hick has further emphasized that since n could in certain circumstances approach 0, eqn. (3) cannot have general validity since $\log n$ would then approach $-\infty$.

Equation (1) and (2) have some intuitive support as well as the theoretical advantage pointed out by Hick : the larger the range of possible signals and responses, the more time the subject is likely to spend scanning them for the occurrence of a signal compared with the time he spends 'wondering when a signal is coming', i.e. in 'scanning' the absence of signal. At the same time it is reasonable to suppose that with practice he might turn his main attention to possible signals or responses and away from the possibility of 'no signal'. The extent to which he did so would be marked by a reduction of the $+1$ in eqn. (1) and its approach to 0. Again, if the subject's attention were positively deflected away from the possible signal sources, a figure in excess of $+1$ might be needed.

What is clear is that eqn. (1) gives a better fit than eqn. (3) to almost all the available data including not only Merkel's and Hick's but also Brown's (1960) and Hyman's own, some of which are shown in Fig. 4. Hyman's experiments required the subject to respond verbally to lights which varied from 1 to 8 in different trials. The regression line in Fig. 4 has been fitted by eye to pass through the origin. It clearly is a very reasonable fit. The best fit would, however, have been obtained by using a quantity a little less than $+1$ in eqn. (1). This is perhaps understandable in view of the fact that Hyman's subjects were given a warning 2 sec before each signal—a length of time short enough for them to have been able to estimate the time of arrival more accurately than was possible in Hick's experiment.

The only results so far found by the present writer to be fitted better by eqn. (3) than by eqn. (1) are from an experiment by Venables (1958) who used a method closely similar to Hyman's in a comparison of schizophrenics and normal subjects. Results from the latter conformed better to eqn. (1), but those from one of the two groups of schizophrenics were more in line with eqn. (3). The other group of schizophrenics produced somewhat scattered results which could be fitted equally well by either equation. It should be noted that both groups of schizophrenics differed from normals in the addition of a substantial constant time for all degrees of choice. Their results would seem, therefore, to need special consideration anyway, which would take us into a detailed discussion of changes occurring in schizophrenia and carry us beyond the scope of this paper.

Whichever side is favoured in this controversy, the general line of approach is supported by three further findings :

(*a*) The amount of information conveyed by the signals in a choice reaction task is reduced if they are not all of equal frequency. The amount of information due to uncertainty about which signal will occur can be worked out by summing the amounts of information conveyed by each signal weighted

according to the probability of its occurrence. Let us call this log n_m—the *equivalent* number of equi-probable choices and corresponding to log n in eqn. (1). We can thus write :

$$\log n_m = \Sigma_i\left(p_i \log \frac{1}{p_i}\right) \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where p_i is the probability of each signal in the set taken in turn. To take account of the requirement to avoid premature or omitted reactions we need to write in place of eqn. (1) :

$$\text{Choice reaction time} = K \log (n_m + n_t) \qquad . \quad . \quad . \quad . \quad . \quad (5)$$

where n_t is the addition due to the need to avoid premature or omitted reactions. This is, in fact, the basic form of eqn. (1) since if all the p_is are equal $n_m = n$ and n_t normally $= +1$.

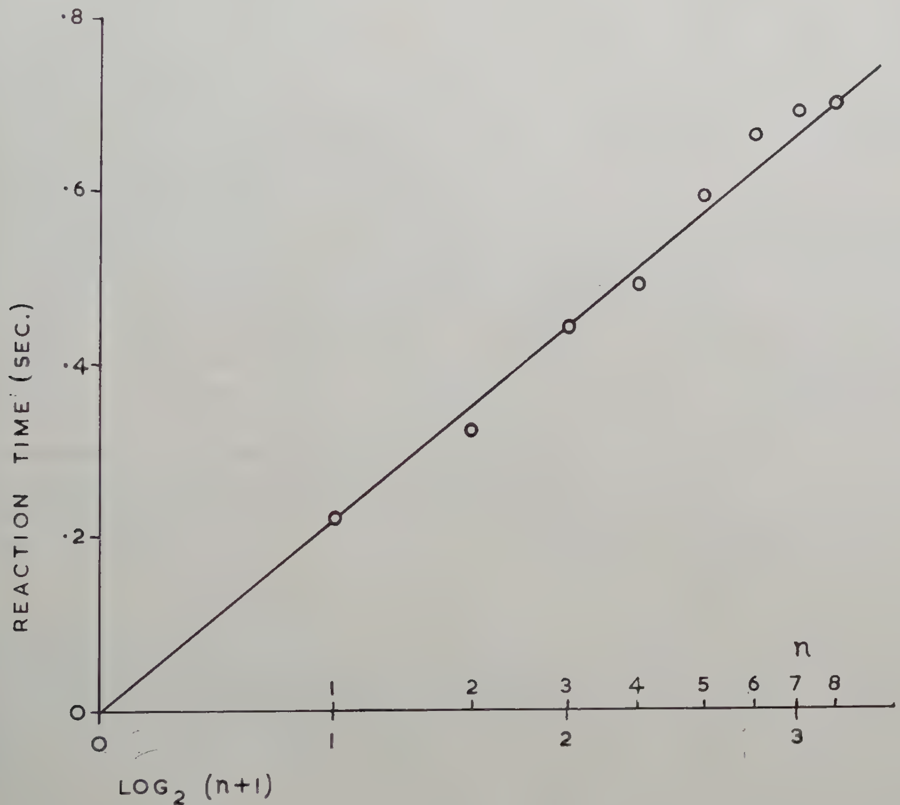

Figure 4. Data from an experiment by Hyman, plotted in terms of eqn. (1). The total number of reactions represented is about 20 000 : 5000 from each of four subjects.

Hyman (1953) and Crossman (1953) both found reductions of about the expected amounts in the average times required to make choices when the signal frequencies were unequal. Hyman's task has already been briefly described. Crossman's was not, strictly speaking, a reaction time task at all, but involved sorting a pack of cards into various different categories such as picture/non-picture. The times for sorting in these ways were compared with the times for sorting into categories of equal frequency such as red/black or the four suits.

(*b*) Reduction of information also occurs when the signals tend to follow each other in recognizable sequences or when any signal is followed by any other more often than would be expected by chance, even though the overall signal frequencies are equal. Hyman (1953) also found the expected shortening of average reaction times in these cases.

(*c*) The amount of information gained is reduced if the subject makes errors. A convenient method of calculating the amount of information gained when errors are present is to make a table with, say, a column for each signal and a row for each response, and to enter the responses made to each signal in the appropriate cells. We can then write :

$$\log n_m = \log T + \frac{1}{T}\sum_{\mathrm{SR}}\left(f_{\mathrm{SR}} \log \frac{f_{\mathrm{SR}}}{f_{\mathrm{S}} \cdot f_{\mathrm{R}}}\right) \quad . \; . \; . \; . \; . \; . \; . \quad (6)$$

where T is the total number of readings, f_{S} is the total frequency of signals in each column taken in turn, f_{R} is the total frequency of responses in each row taken in turn, and f_{SR} is the frequency of readings in each cell taken individually. Hick (1952 a) sets out a practical example from his experiments.

Hick found that the shortening of reaction-times when substantial numbers of errors were made was by approximately the amounts expected. This finding is important because it gives us a rational means of combining speed and accuracy of performance into the single score of *amount of information gained*. It also means that times for different tasks can be regarded as comparable only if errors are held constant, and conversely that error rates can be properly compared only if times are held constant.

The same considerations imply that the $+1$ in eqn. (1) will be reduced if premature responses occur or if responses are omitted.

Hick has noted in a private communication that most, perhaps all, experiments in this area have tacitly assumed that all errors are equally ' bad ', and suggests that this may not in fact be correct. Further work would seem to be required on this matter.

2.3.1. *Conceptual ' models '*

Although the logarithmic relationship between choice reaction time and degree of choice has now been found in a substantial number of experiments, it is by no means clear what the subject is doing when making a choice. We may consider the problem under two heads : firstly the relative roles of *perception* and *selection of response;* secondly the detailed methods of making choices.

A. *Location of choice within the chain of central mechanisms*

The important evidence on this question is the difference of slope between the symbolic and non-symbolic tasks in Crossman's (1956 a) experiment shown in Fig. 3. The slope of the symbolic task is very similar to that in Merkel's, Hick's, Hyman's and Venables' experiments, all of which involved an arbitrary or unfamiliar translation from signal to response. Crossman also lists a number of other experiments where a similar rate of 5–8 ' bits ' per sec was found although the times for various degrees of choice were not systematically explored. Crossman's non-symbolic task yielded an incremental rate of

between two and three times this—a rate similar to that found for reading aloud a list of words or nonsense syllables arranged in random order (Hick 1952 b).

The evidence suggests that the difference of slope is due to the greater involvement of the translation mechanism in the symbolic than in the non-symbolic case. If the relationship between signals and responses is simple and straightforward, as in Crossman's case, or if it is extremely well learnt, as with naming of letters of the alphabet, it is reasonable to suppose that all ' connections ' in the translation mechanism can be pre-set and held ready-made so that the subject's response will be ready as soon as the signal has been identified. If so, the rate for non-symbolic tasks gives a measure of the speed of operation of the perceptual mechanism. It is uncertain whether the rate for symbolic tasks reflects the action of both the perceptual and the translation mechanisms, or whether their activities overlap so that the overall rate is that of the slower. Crossman (1955) has found evidence that discrimination of signal and choice of response can overlap, and further evidence that this occurs in some cases will be presented later. We may, therefore, tentatively assume that the overall rate is that of the slower process—that is of the choice of response by the translation mechanism.

The role of the translation mechanism is, perhaps, also indicated by the fact that the time taken to sort playing cards by number is substantially shorter if the piles are arranged in normal numerical order than if they are in random order. In the former case the subject has a well-established ' set ' relating numbers on the card to the appropriate piles, so that most of the ' work ' of translating from signal to response is cut out. The effect has been well shown in an experiment by Morin and Grant (1955) in which the relationships between the positions of signal lights and response keys were varied systematically from completely straightforward to random.

Further indications of the part played by the translation mechanism are given by the lengthening of reaction times when each response is attached to two or more signals instead of to only one (Crossman 1953, Griew 1958 b, Rabbitt 1959), or when signals consist of different combinations of a few display lights instead of each being a single light (Gibbs 1952; Brown 1960).

What appear at first sight to be extreme cases of pre-setting the translation mechanism come from two experiments in which the relationships between signal and response were deliberately made very straightforward and the subjects given unusually long practice. The first of these was by Mowbray and Rhoades (1959). Responses were made by pressing buttons placed in convenient positions under the subject's fingers, and the display lights were on a panel in the same relative positions : no differences of mean reaction time were found between two and four choices. The second experiment by Leonard (1959) went even further in relating signals to responses. The latter were made by pressing the armatures of relays placed conveniently under the subject's fingers, and the signals consisted of vibrations in these same armatures. Leonard found a rise of reaction time from simple to two-choice, but none thereafter to four- and eight-choice.

The view that, with sufficient practice and a straightforward relationship between signal and response, the translation mechanism can be pre-set to such an extent that it causes *no* delay beyond two choices suggests an important

effect of practice and opens the way to a stimulating line of research. It is however speculative, and a very much simpler explanation for these two experiments will be outlined later.

B. *Detailed activities in making choices*

Several ' models ' implying different processes of making choices have been proposed but none appears so far to be entirely satisfactory. They can, however, act as valuable stimulants to thinking. We shall examine four in turn :

(i) *Serial dichotomous classification.* The most readily conceived model is one of those examined by Hick (1952 a). The subject is thought of as making a series of sub-decisions each taking the same time. With the first he identifies the signal and the response to it as lying within one-half of the total possibilities ; with the second as lying within one-half of this half, and so on until the specific signal and response have been found. He is not able to make his divisions into exact halves unless N is an exact power of 2, but can approximate closely to doing to. This model conceives of errors as being due to the subject not carrying the process of elimination far enough, thereby saving one or more sub-decisions and the time these would take, but risking error because the final choice is made ' at random ' among the possibilities that remain. The fact that a subject might not—indeed cannot always—reject exactly half the possibilities remaining at any sub-decision is not a serious criticism of the model since it would still give approximately the correct result so long as the subject started by rejecting broad classes of possibility and then went on to reject finer classes within a broad class chosen. What is a serious criticism is that the model would not predict any reduction of reaction time when the frequencies of different signals were unequal. For example, when there are two signals, would require one sub-decision to be made in all cases whether or not the signal frequencies were equal.

(ii) *Serial classification with check on accuracy.* The difficulty about unequal frequencies can be overcome by assuming that if during a sub-decision the subject finds that the signal is *not* in the half he is examining, he checks that it *is* in the other half before proceeding. According to this model each dichotomous sub-decision might require either one or two inspections and, if each of these takes time, the sub-decision will be completed in a shorter or longer time according to whether or not the first inspection is successful. We might assume that a subject would, if signal frequencies were not equal, tend to try the more likely signal first. Doing so would mean that he would save time since he would thereby increase the chance of his first inspection being successful.

If we assume that the times are equal to make an original inspection and to make a second, check inspection if the original inspection is wrong, the average time required for a binary decision where signal frequencies were equal would be made up of 50 per cent cases where one inspection-time was sufficient and 50 per cent where two were needed. The mean time would be the mean of 1 and 2, i.e. 1·5.

If one alternative occurred three times as frequently as the other, and the subject biased his first choices to correspond to the signal frequency, trying the more frequent alternative first three times out of four and the less frequent once out of four, he would identify the more frequent signal in one inspection

and the less frequent in two, three times out of four. On the remaining one out of four occasions the number of inspections would be reversed. The average number required works out at 1·375. If the subject *always* tried the more frequent alternative first the average would be further reduced to 1·25.

The effect of frequency unbalance is limited by the fact that, if our assumption is correct, the ratio of the times required to respond to the more and less frequent alternatives can never exceed 1 : 2. Two points follow from this, both of which are in line with Hyman's results. Firstly, the times will tend to be longer for more frequent signals and shorter for less frequent than a strict information analysis would predict. Hyman specifically notes the occurrence of this fact in his paper. Secondly, the effects of extreme frequency unbalance will be less than those predicted on information analysis. Close scrutiny of Hyman's results indicates evidence that this was so in his experiment : for example, the reaction times for two choice tasks with signals presented in the ratios 1 : 9 and 1 : 4 were almost identical.

Two further implications of this model should be noted. Firstly, it opens up a new approach to the concept of ' expectation ' as a factor shortening reaction time by suggesting that reaction times may be appreciably affected by the order in which inspections are made, and that this order can be partially determined by the subject before the signal arrives. In other words, he can decide beforehand which of the choices open to him he will try first. If his 'expectation' is correct more often than not, he will reduce the number of inspections required when the signal actually arrives, and thus shorten reaction time. When averaged over a large number of readings the effects of such pre-determination of choice might well suggest that expectation should be conceived as a continuous variable, although it would in fact be essentially discontinuous when examined in detail.

The second implication of this model is that the strategy of making successive binary decisions is not always the most efficient : for choices between two and eight, it is better to make a serial examination of each possibility in turn. This procedure requires one inspection for each possibility examined up to and including the correct one, so that a subject will, on average, arrive at the correct choice in $(N+1)/2$ inspections. This number is shown in Table 1 to be less than the number required by methods employing successive dichotomization. Beyond eight choices the method of serial examination is inferior to dichotomization.

Further economies can be achieved if, for choices of six and beyond, the subject divides the possibilities into threes or fours and explores these serially

Table 1. Average numbers of inspections required using different strategies to decide among $n+1$ alternatives

n	$\log_2(n+1)$	$1{\cdot}5 \log_2(n+1)$	Strict dichotomizing	Running along line i.e. $(N+1)/2$	Optimum strategy
1	1·00	1·50	1·50	1·5	1·00
2	1·59	2·38	2·44	2·0	1·50
3	2·00	3·00	3·00	2·5	2·00
4	2·32	3·48	3·55	3·0	2·50
5	2·59	3·89	3·94	3·5	3·00
6	2·81	4·22	4·25	4·0	3·50
7	3·00	4·50	4·50	4·5	3·71
8	3·17	4·75	4·80	5·0	3·75

rather than exploring all signals serially or following a strict dichotimizing strategy. There would also be a substantial saving of time if the subject could so bias his inspection of the display before the signal arrived that he never scanned for absence of signal but always for a positive signal. Such a bias could account for the well-known tendency for reaction to be faster if the subject attends to the response he is about to make than if he attends to the signal he is expecting to receive (see Woodworth 1938, Woodworth and Schlosberg 1954). In the latter case he is likely to be checking all the time whether or not the signal has arrived. In the former case he is 'set' to 'signal only'. The observed difference of simple reaction time between these conditions (a ratio of about 1/1·5) appears, on average, to be of the magnitude expected although there are wide differences between different experiments. The greater scatter of simple reaction times found when the subject attends to the signal is also in line with the present model, since the times would consist of some in which two inspections had been required and some in which one had been sufficient. If he had attended to the response only one inspection would have been needed on all occasions.

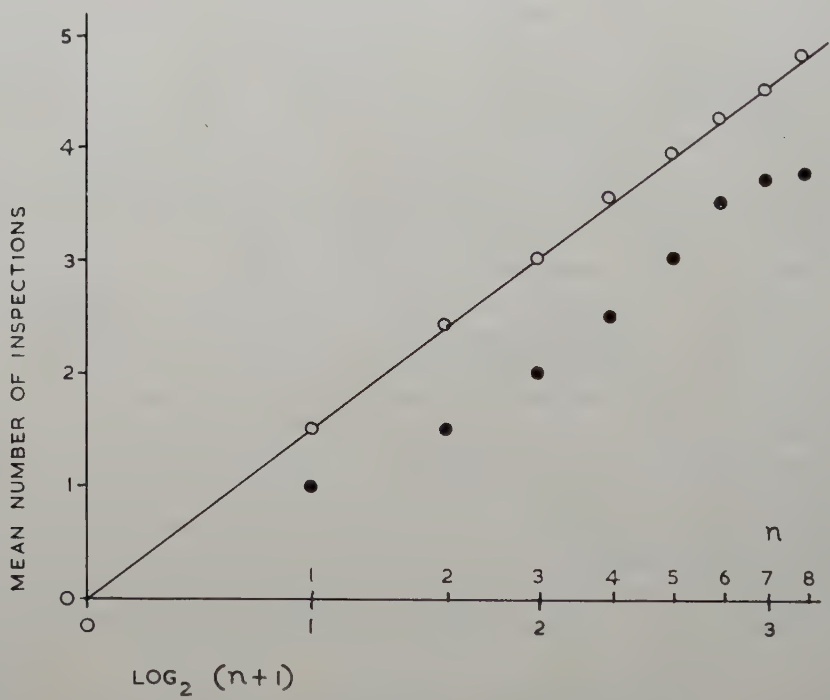

Figure 5. Average numbers of inspections required using different strategies. The continuous line is $1{\cdot}5 \times \log_2(n+1)$, the open circles are for the strictest possible dichotomizing strategy, and the filled circles are for the optimum strategy according to Model (ii).

The joint effects of these various economies are shown in Table 1. and Fig. 5 The curve looks very much like an exaggeration of the slightly S-shaped curve obtained by Hyman (Fig. 4). We may, perhaps, suppose that Hyman's subjects managed to approximate to this strategy without achieving it completely.

The fact that this shape is not reflected in Merkel's or Hick's data is not a crucial argument against the present model. The attainment of maximum efficiency is likely to be rare. It involves the maintenance of a strategy in the face of many temptations to vary it by guessing at what signal is likely to arrive.

The precise strategy adopted is likely to vary with the way in which signals are arranged and with the relationship between signal and response. If there are very large numbers of, say, lights and keys in a display and they are not split up into groups, counting from one end may be the only method of ensuring an accurate response. Where the relationship between signal and response is symbolic, it may well be that dichotomizing is favoured, but where signals and responses are both laid out in a row, or in the same pattern, serial exploration may be preferred. If so, the model proposed here suggests a rather simple explanation of Mowbray and Rhoade's and Leonard's experiments when the responses are considered in detail.

Table 2. Fingers used for responses in experiments by Mowbray and Rhoades (1959) and Leonard (1959)

Number of choices	1	2	4	8
Mowbray and Rhoades	—	Right and left index	Right and left index and middle	—
Leonard				
Group I	Right index	Right index and middle	Right all fingers (not thumb)	Right and left all fingers (not thumbs)
Group II	Right index	Right and left index	Right and left index and ring	Right and left all fingers (not thumbs)

The fingers used for different numbers of responses are shown in Table 2. It is important to note, however, that results were in each case analysed for only one finger, the left index by Mowbray and Rhoades and the right index by Leonard. Now Table 1 shows that with the present model substantial savings can be obtained by serial exploration or by mixing serial and dichotomizing strategies instead of using a pure dichotomizing procedure. Some rise of time from simple to two choices seems inevitable, but for four and eight choices the subject will do best if he starts at one end of a line and inspects in turn until he finds the signal and response required. It seems probable that he would tend to start by inspecting one of the two signals and responses used in the two-choice situation, and if he does this he is likely to achieve very similar times with the index fingers under all conditions. The correctness of this explanation could be tested by analysing the response times for all fingers used and by arranging the experiment so that all possible combinations of fingers were used in different trials for each degree of choice. Results which fail to do this need to be taken with extreme caution because of Crossman's (1956 a) finding that in his non-symbolic task differences of reaction time for different positions in the row far exceeded those for different degrees of choice.

(iii) *Simultaneous scanning.* Hick suggested as an alternative model that the subject might compare an incoming signal simultaneously with all possible identifications. This model has been stated explicitly by Christie and Luce (1956) and discussed by Rapoport (1959). It postulates that the comparison of the incoming signal with each possible identification takes a time which is

subject to random variation and that the decision to act is only taken when all identifications have been completed. The larger the number of possible identifications, the greater is the likelihood of one taking a long time and thus causing a long reaction time.

This model has the advantage that it assumes a continuous distribution of reaction times. It is not easy to see, however, how it accounts for the effects of unbalanced frequencies, and the variances it postulates do not fit well with those found by Hick (1952 a).

(iv) *Neurological models.* The three models discussed so far have been in terms of what may be called the *micro-behaviour* of the subject, and have attempted to break down the decision process into a series of subsidiary actions. The building of *neurological* models would seem to be premature in the present state of knowledge. However, two examples deserve brief mention.

(*a*) The first of these is the discrimination model of Householder and Landahl (1945) which postulates that discrimination is achieved by a ' nerve net ' in which each signal tends to excite its corresponding response and to inhibit the excitatory tendencies of other signals. Similarly the excitation of any response tends to suppress other conflicting responses. The strongest excitatory tendency tends to capture the response mechanism and lead to action. The fact that this model contains several parameters reduces its immediate usefulness, although the authors make suggestions about what further knowledge is required to determine the parameters concerned.

(*b*) A less detailed model recognizes that any signal or response will be associated with activity in a restricted set of nerve cells, and that for it to be distinguished from other signals or responses the level of activity in these cells must exceed by a given amount that in the cells responsible for other identifications or responses. It further assumes that the initial effects of an incoming signal will be relatively diffuse, and that the building up of more specific activity will take time. Decision is assumed to be completed when activity has become intense enough in the region of one response for it to be sufficiently *differentiated* from the adjacent responses. If the specificity of response is low—as, for example, if the response can be made with the whole hand, the build-up need not be large. If, however, a response has to be made with one finger and not with adjacent fingers the build-up will have to be very much greater and thus take more time. We may, perhaps envisage that the build-up results in part from the incoming signal, but very largely from self-regenerating circuits in the brain. These latter would tend to build up activity exponentially with time thus producing the kind of relationship of eqn. (2) with N conceived not as mere number of alternatives, but as the number of alternatives to be distinguished within a given area of brain tissue, the amount of tissue devoted to each diminishing as N increases.

The model can deal with errors by assuming either that a decision is made too soon, or that a ' burst ' of random neural activity upsets the pattern at the crucial moment. It can deal with the effects of unequal frequencies and of expectation by assuming that the subject activates in advance the tissue corresponding to certain choices so that when the signal arrives the build-up required to produce a decision in favour of these is less than for others.

2.4. *Control of Movement*

We have seen that the experiments on choice give a tentative measure of the capacity of the translation mechanism. The next portion of the chain of mechanisms to be tackled was the central control of movement.

The relation between amplitude, accuracy and the time taken to make hand movements had been a subject of discussion for some years before Fitts (1954) suggested a formulation in information-theory terms which connected all three together. Fitts proposed essentially that

$$\text{Movement time} = a + b \log (2A/W) \qquad (7)$$

where W is the *width* of the 'target' within which the movement is required to end, measured parallel to the direction of the movement ; A is the amplitude of the movement measured from its starting point to the centre of the target ; and a and b are constants. The essential point of this formulation is that it makes movement time constant for any given ratio between amplitude and target width.

Fitts recognized that the multiplication of A by 2 was arbitrary although some such procedure appeared necessary in order to ensure that the logarithm was always a positive quantity : had the fraction A/W been used, W would have exceeded A in the case of movements made from just outside a target and this would have made log (A/W) negative. There was also some conceptual plausibility in that the subject could be thought of as having to 'choose' a movement which had the possibility of either over- or under-shooting the target. Fitts also recognized that it would not always be accurate to take W as the measure of the scatter of the shots : the subject might, for example, concentrate his shots in a narrower width, in which case the movement time would be expected to be greater than that predicted by eqn. (7).

Fitts backed his formulation with four sets of experimental data. In one experiment the subject was required to 'dot' alternately on two metal strips 6 in. long using a metal-tipped stylus weighing 1 oz. and about the size of a pencil. The long axes of the strips were perpendicular to the line of movement between them. The strips were mounted on a board placed in such a position that the subject's movements were from side to side in front of him. Four widths of target strip were used : 2, 1, 0·5 and 0·25 in. at each of four distances between centres : 2, 4, 8, and 16 in.

Average times per movement, plotted according to eqn. (7) with A taken as the distance between target centres, and W as the width of the target strips, are shown in Fig. 6. All the points, except those at the extreme lower end, lie close to a straight line and suggest that Fitts's formulation was right in principle. Fig. 6 does, however, have three unsatisfactory features which suggest that some modifications of detail are required :

(*a*) The straight line running through most of the points cuts the zero information line below the origin making a in eqn. (7) a negative quantity. Crossman (1957) has suggested that this difficulty can be avoided by omitting to multiply A by 2 so that eqn. (7) becomes

$$\text{Movement time} = a + b \log (A/W) \qquad (8)$$

The constant a works out at about 0·05 sec, a figure close to that found by Crossman himself in a similar experiment for the time spent on the targets as opposed to the time spent moving between them. It seems obvious at first

sight that the time spent stationary on the targets ought not to be included in the measure of movement time. It has been found, however, in another similar experiment (Welford 1958) that results are, on the whole, more uniform in their fit to eqn. (7) or (8) if the times on target and between targets are added together than if the latter measure only is taken. It seems as if there may be some compensatory tendency at work which makes it possible for time spent on the targets to be used to shorten time spent moving, at least to some extent. Such a compensatory effect would be consistent with Fitts's view that movement time is limited by the speed at which central processes can control and monitor movement. Such processes must obviously to some extent precede and outlast the movements concerned, and the time thus required would be spent on the targets.

(*b*) Although the points in Fig. 6, except those at the extreme lower end, lie fairly close to a straight line, the best fitting line through them would curve

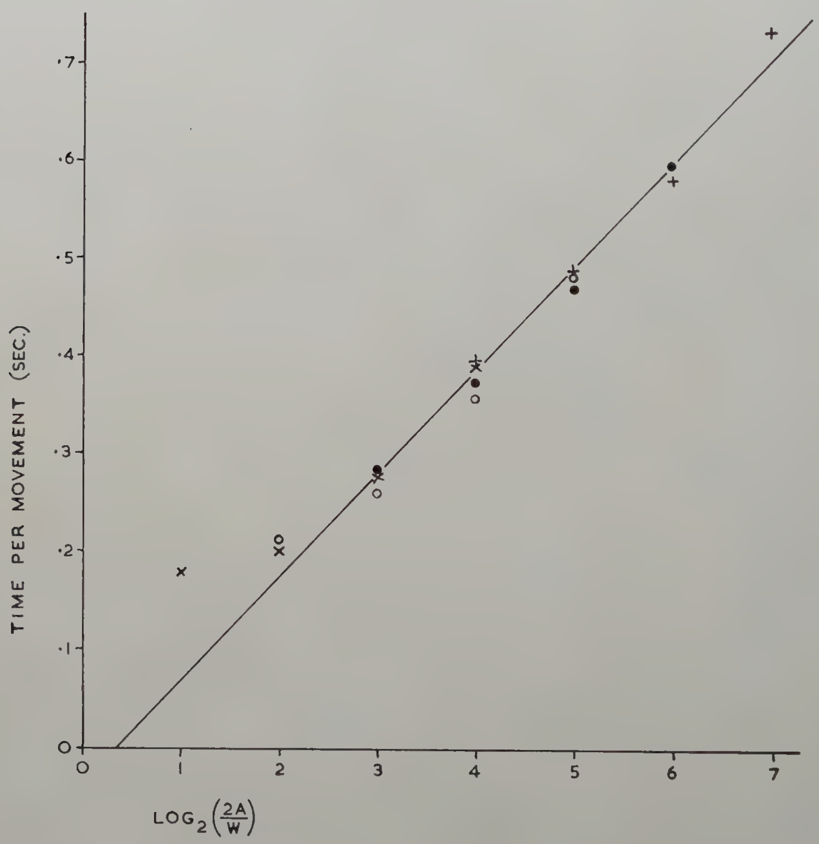

Figure 6. Times for reciprocal tapping with a 1 oz. stylus plotted in terms of eqn. (7). Data from an experiment by Fitts. The target widths (W) are indicated as follows : $\times = 2$ in., ○ = 1 in., ● = $\frac{1}{2}$ in. and $+ = \frac{1}{4}$ in. The four movement amplitudes (A) for each target width are, from left to right, 2, 4, 8 and 16 in. Each point is based on a total of 613–2669 movements obtained from 16 subjects.

gently upwards. The curve can be substantially removed by making a further modification to the equations by writing

$$\text{Movement time} = K \log\left(\frac{A+\frac{1}{2}W}{W}\right) = K \log\left(\frac{A}{W}+0{\cdot}5\right). \qquad (9)$$

This formulation makes movement time dependent upon a kind of Weber fraction in that the subject is called upon to distinguish between the distances to the far and the near edges of the target. To put it another way, he is called upon to choose a distance W out of a total distance extending from his starting point to the far edge of the target. The formulation also preserves the advantage which Fitts claimed for the procedure of multiplying A by 2, in that the logarithm can never be negative, since in the extreme case when the movement begins at the edge of the target $A = \frac{1}{2}W$.

(*c*) The curve in Fig. 6 shows a distinct flattening at the lower end. This is probably due, as Crossman has suggested, to some limiting factor setting a minimum time per movement however short or unconstrained. Study by the present author of dotting between targets using a pencil as a stylus suggests that this limiting factor affects the amount of target used. When the targets are wide and the distance short the subject uses very much less than the full target width. He is, in fact, transmitting more information than a calculation in terms of eqn. (9) would assume because the effective W is narrower. The narrowing of W is to some extent reflected in a reduction of errors and if due allowance is made for them eqn. (9) still holds reasonably well.

The method of correcting for errors has been described by Crossman (1957). It makes use of the fact that the information in a normal distribution is $\log_2 \sigma\sqrt{(2\pi e)}$, where σ is the standard deviation of the distribution. The shots on the targets in ‘ dotting ’ tasks such as that of Fitts do appear to fall into an approximately normal distribution. Now $\sqrt{(2\pi e)} = 4{\cdot}133$ and a range of $\pm$ half this, i.e. $2{\cdot}062\sigma$, includes about 96 per cent of a normal distribution. We can therefore argue that if about 4 per cent of shots fall outside the target, $\log_2 W$ is an accurate representation of the information contained in the distribution of shots. We can also argue that if the errors exceed 4 per cent the *effective* target width is greater than W, and if the errors are less than 4 per cent the effective target width is less than W. How much greater or less can be calculated from tables of the normal distribution (e.g. Fisher and Yates 1957, Table 1). For example, suppose $W = 2$ in. and the errors are 1 per cent. Then the effective $W = 2 \times 4{\cdot}133/5{\cdot}152 = 1{\cdot}604$ in., since all but 1 per cent of a normal distribution lie within a range of $\pm 2{\cdot}576\sigma$ (i.e. $\frac{1}{2} \times 5{\cdot}152$) of the mean.

It has been assumed so far that errors are distributed equally to both sides of the target and that the mean is in the centre of the target so that the mean amplitude of movement is correct. If the mean falls to one side or the other of the centre, the effective A will be different from the distance A between target centres. Strictly speaking, the effective A should in any case be the geometric mean of the distances between each shot and the next, although it is usually approximately correct to take the distance between the means of the two distributions of shots. If the distributions can be assumed to be symmetrical it is not necessary to record the positions of all the shots in order to calculate the positions of the means : the mean of each distribution can be

inferred from the proportions of shots falling on the near and far sides of the target concerned.

Fitts's data are plotted in Fig. 7 using eqn. (9) and making appropriate adjustments to W for errors. The data were unfortunately not given which would have made it possible to calculate the effective A's, and these have been assumed to be correct. The results lie close to a straight line which passes through the origin.

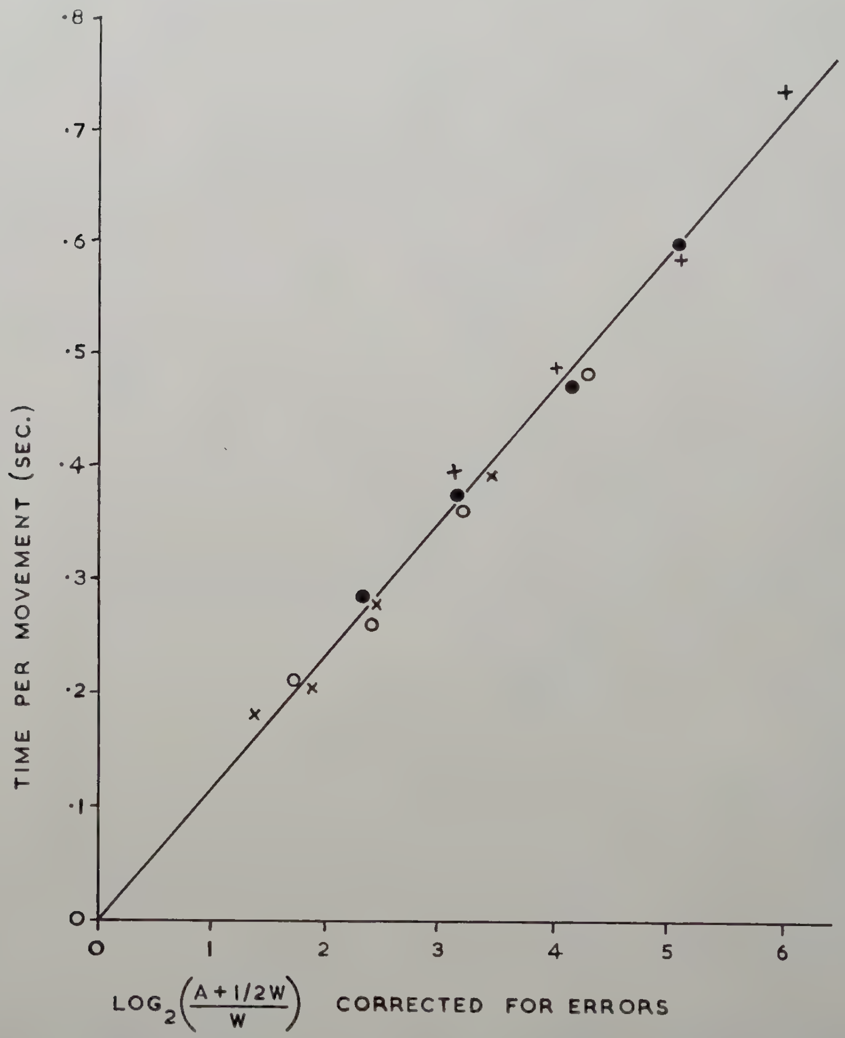

Figure 7. The same data as in Fig. 6 plotted in terms of eqn. (9) corrected for errors by Crossman's method. Target widths and movement amplitudes are indicated as before.

2.4.1. *Further data on movement times*

Crossman (1957) has produced results which although obtained independently of Fitts's, are in striking agreement with them. He has also shown that the length of the target strip as well as its width affect movement time although to a smaller extent. The reason for this becomes clear when

one studies records of dotting with a pencil between targets drawn on paper. The shots in each target fall roughly into an ellipse with its long axis parallel to the line of movement : the scatter perpendicular to the line of movement is very much smaller than that parallel to it.

Closely similar results to those of Figs. 6 and 7 were obtained by Fitts for a similar ' dotting ' task but using a stylus weighing 1 lb instead of 1 oz. Fitts notes that this similarity clearly supports the view that movement time is determined more by the central processes controlling and monitoring movement than by any factors of muscular effort involved.

In a third experiment by Fitts the subject was required to transfer discs with holes in their centres from a vertical pin to another similar pin a given distance to the left. Four distances between pin centres were used with each of four sizes of hole. The ' target widths ' were taken as the differences between the diameters of the pins and of the holes.

Fitts's fourth task was the transfer of metal pins from one set of holes to another a given distance away. Four diameters of pins were used with each of five distances. The holes were in each case twice the diameter of the pins : the ' target width ' was again taken as the tolerance between the pins and the holes.

The results of these last two experiments are, unfortunately, not suitable for detailed examination as the times taken to transport the discs or pins to their ' targets ' were not measured separately from the return (' transport empty ') movements of the hand to pick up the next disc or pin. We cannot assume these return movements to be of constant duration, or of a duration proportional only to distance. Evidence that they *may* be comes from results by Annett *et al.* (1958), but Crossman (1957) found that return movements tended to become slower as the accuracy of the outward movements rose. There is considerable support for the view that the speed of one movement in a cycle tends to affect those of others (de Montpellier 1935; Wehrkamp and Smith 1952; Denton 1953; Welford 1958, p. 105; Simon and Simon 1959).

The difficulty does not attach, however, to the pin-transfer experiment of Annett *et al.* (1958). Their subjects, like Fitts's, were required to transfer pins from holders to sockets. The pins were $\frac{1}{8}$ in. in diameter and different sets of sockets ranged in diameter from 9/64 through 12/64 and 24/64 to 72/64 in., giving tolerances of 1/64, 1/16, $\frac{1}{4}$ and 1 in. respectively. The distance between centres of holders and sockets was in all cases 8 in. The holders were 9/64 in. diameter.

Movements were recorded by a film taken at 48 frames per second, and from this it was possible to determine the interval elapsing between the pin leaving the holder and its release after having been placed in the socket (' movement loaded ' plus ' position '). This interval was the time taken to make a movement the accuracy of which could be defined in information terms. ' Target width ' was taken as the difference of diameter between pin and socket.

The average times taken by the three subjects are plotted in Fig. 8 in accordance with eqn. (9). The points in this lie close to a straight line, but the times for the finer tolerances tend to be too short. The line would have been straight and have passed through the origin if the pins had been 7/64 in.

in diameter instead of $\frac{1}{8}$ in., and this fact led the present writer to consider whether there were any reasons why the *effective* pin diameter might have been a little less than $\frac{1}{8}$ in. Two factors appeared possible : firstly that the ends of the pins were slightly rounded, and secondly that the subjects were applying them to the holes at an angle so as to present a rounded edge which would 'find' the hole more easily than would the flat end of the pin. Enquiry from one of the authors revealed that both factors were present : the ends of the pins were slightly chamfered since if this was not done "the subjects often fumbled when putting them into the smallest holes"; and the film clearly showed the pins being applied to the holes at an angle.

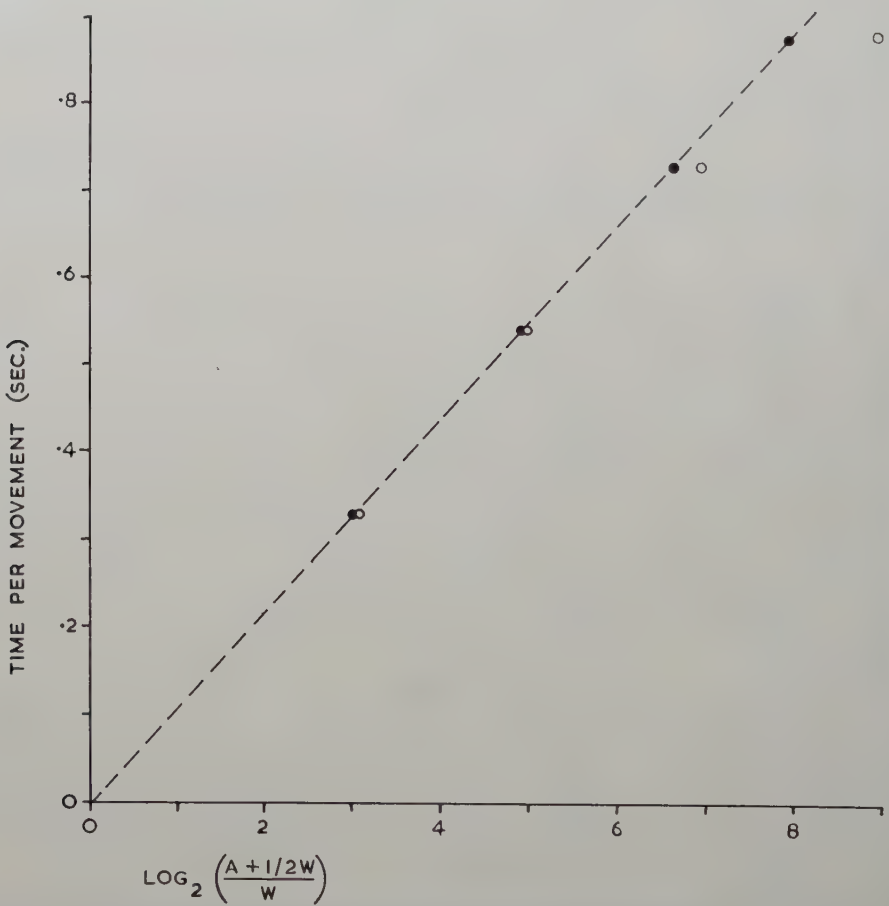

Figure 8. Data from an experiment on pin-transfer by Annett *et al.* plotted in terms of eqn. (9). The open circles indicate the values calculated for the actual pin diameter of $\frac{1}{8}$ in. The filled circles indicate what the values would have been for a pin diameter of 7/64 in. Each point is based on 960 readings : 320 from each of three subjects. Recording began after extensive practice.

Further work by Schouten *et al.* (1960) has indicated that it is necessary in pin transfer tasks to take account of the absolute size of the pin as well as the tolerance between pin and hole, since the finer pins take a disproportionately long time to locate. This is perhaps in line with Crossman's (1957) finding that with small targets both width and height need to be taken into account : small diameter holes (which tend in these experiments to be associated with

small diameter pins) would produce a target which was severely restricted not only in the line of movement but also perpendicular to it.

2.4.2. *Conceptual 'models'*

We have spoken of the subject 'choosing' a target width out of a longer extent but this is only a figurative description. The rates of information gain attained—8–12 bits per second—suggest that performance is not limited by the same mechanisms as in the case of choice-reaction times but, presumably, by the central mechanisms phasing and coordinating movement.

As a detailed model, Fitts suggested that the subject could be conceived as trying to make a series of movements each of uniform extent but subject to random disturbances which might either add to or subtract from the extent aimed at. This model leads naturally, as Crossman (1957) has pointed out, to eqn. (8) since the extent aimed at is equal to A and the random variation, if accuracy is held constant, can be represented by W.

In criticism of this view it seems fair to make two points. Firstly, the subject's task is not, strictly speaking, to make movements of uniform extent, but ones which lie within limits. Secondly, the subject's actions in 'homing' upon a target are more complex than merely trying to make a movement of a given extent (or duration). Annett *et al.* noted in their experiment that most of the variation in time taken with different sizes of hole occurred during the last two inches of movement. Plots of movements suggest that the subject starts a movement fast and gradually slows down as he approaches the target.

Taking these two points together, it is reasonable to think of the subject as having to consider all movements of an extent short of the far edge of the target and to reject all those which are short of the near edge. He starts by rejecting a large class of those which are much too short and then rejects finer and finer classes as the target is approached. His behaviour may be conceived as analogous to that envisaged in the first of the models discussed in relation to choice reactions : approximately equal times might be taken to cover the first half, the next quarter, the next eighth and so on of the total distance to the far edge of the target. For this way of looking at the control of movement the appropriate formulation seems to be eqn. (9).

2.5. *Perceptual Identification and Discrimination*

The work outlined on choice-times has assumed that these include both the identification of the signal and the selection of the response. We have seen that there is some evidence to show that the former takes place more quickly than the latter, but obeys the same logarithmic law. We do not know for certain whether this law applies to the identification of objects in more complex perception, but the approach appears worth exploring.

We may think of a subject called upon to identify an object as having a large, although finite, number of categories 'at risk' into which he might place it. The number might be cut down by instructions or previous knowledge which told him that the object must belong to a certain broad class —say of animals as opposed to plants or inanimate objects—so that when the object is actually perceived, choice has to be made between relatively few alternatives. It is stimulating to think of perceptual identification as proceeding by a series of choices, first into broad categories, later into progressively narrower and more precise ones. Such a view would suggest

that errors of identification could be classed into two types : firstly, cases where the object presented had been placed in a *wrong* category ; secondly, cases where it had been placed in a category which was correct as far as it went but was not precise enough.

The main evidence of the time taken by perceptual processes comes at present from a different type of experimental task. In the study of choice times, the discriminability of signals has been made very clear and the degree of choice varied. We now turn to experiments in which degree of choice has been held constant and the discriminablity of the signals has been varied. Some of the evidence has been known for a long time but formulae were not proposed until recently. It must be said at once that interpretation of results in this field is uncertain and there are some disagreements which clearly call for further work.

We shall base our discussions upon five sets of experimental results :

(*a*) Henmon (1906) exposed pairs of lines on a screen and the subjects were required to indicate on one of two morse keys which of the pair was the longer. The shorter line was always 10 cm and the longer varied in steps of 0·5 cm from 10·5 to 13·0 cm.

(*b*) The same author in another experiment required subjects to discriminate between tones of different frequencies produced by tuning forks. The discriminations required were 4, 8, 12 or 16, cycles from a standard of 499·75 per second. The task was to decide whether the variable presented at any trial was higher or lower than the standard.

(*c*) Crossman (1955) required subjects to sort specially prepared packs of cards according to numbers of spots on the cards. Each pack contained equal numbers of cards with each of two different numbers of spots : 10/1, 10/5, 12/8, 12/9, 10/8 and 12/10. Crossman also conducted further experiments, including one with cards in which each pack contained cards with three, four or five different numbers of spots, and one in which small metal canisters were sorted by weight. All the results agreed closely and discussion will therefore be confined to the one experiment since any treatment satisfactory for it is likely to be suitable also for the others.

(*d*) Birren and Botwinick (1955), in an experiment somewhat similar to Henmon's, exposed pairs of lines, and the subject said ‘ left ’ or ‘ right ’ to indicate which was the longer. Reaction times were recorded by means of a voice key. The lines were exposed for 2 sec, a time longer than most of the reaction times. One line was always 80 mm long and the other was from 1 to 50 per cent shorter.

(*e*) Botwinick *et al.* (1958) repeated this last experiment using two different exposure times : 2 sec and 0·15 sec.

Data by Lemmon (1927) were examined but it was found, in agreement with Crossman (1955), that the experiment was conducted in such a way that the results were not suitable for detailed scrutiny.

All these experiments show that times tend to rise as discrimination becomes finer. The times are affected to only a small extent by the absolute magnitudes of the signals but are essentially a function of the ratio of one to another. Any formula or law relating reaction time to fineness of discrimination must therefore provide for time to rise as the ratio between the quantities concerned becomes smaller but to be little affected by changes in absolute magnitude, except perhaps at very low values.

Crossman (1955) considered three types of formulation in relation to Henmon's data and his own.

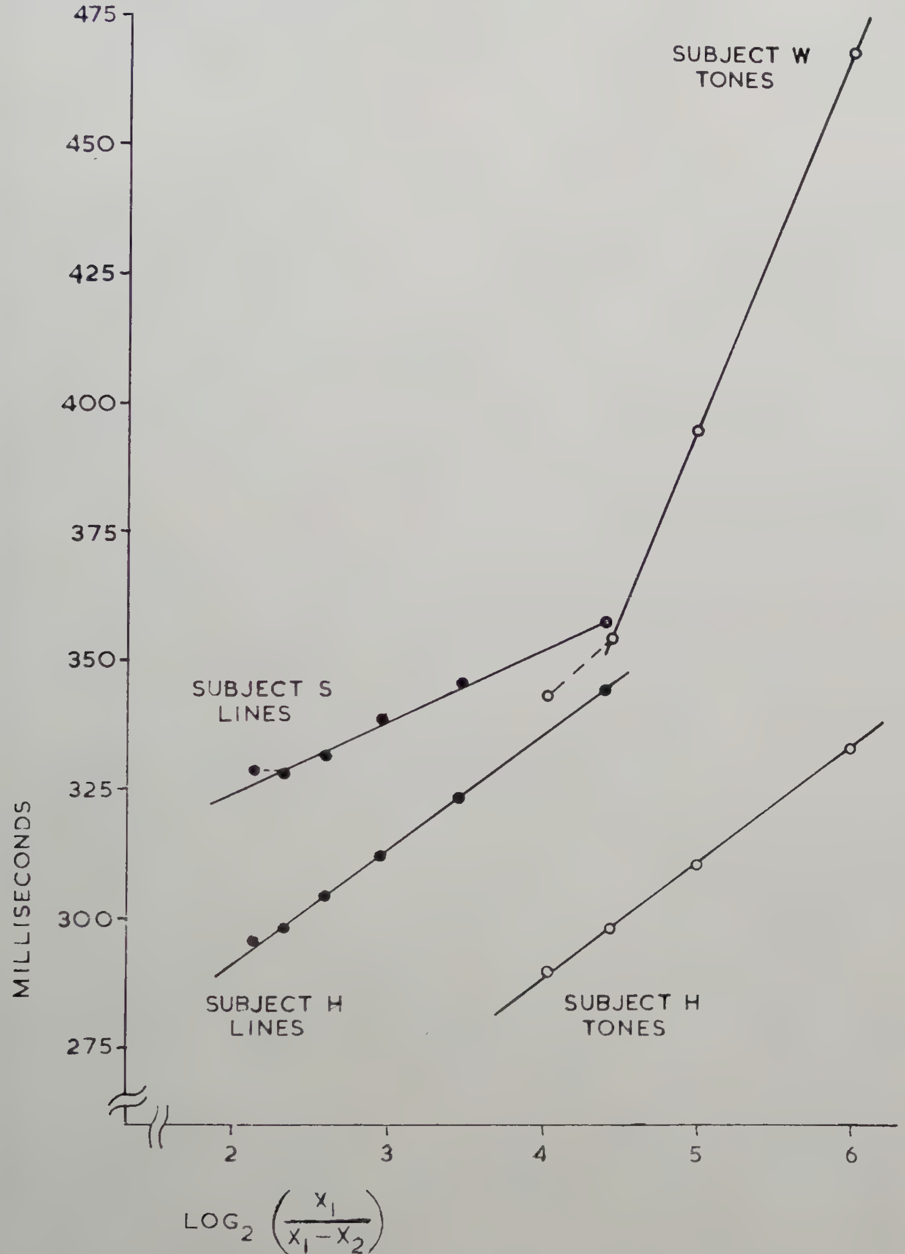

Figure 9. Henmon's data plotted in terms of eqn. (10). Each point for discrimination between lines is based on 400 readings, and for discrimination between tones on 320 readings. In the experiment with tones the subjects may either have made their discriminations between the variable and the standard at each presentation, or have taken other trials into account and thus compared the two values of the variable on opposite sides of the standard. The latter procedure would enable discrimination to be made of a frequency difference double that of the former. The points in the Figure are plotted on the assumption that this second procedure was followed. If it is assumed that the first procedure was used, all the points for tones should be moved one " bit " to the right.

(i) *An information-theory model.* This was based on Wiener's (1948, p. 75) formula for measurement and makes the time taken by discrimination a linear function of the logarithm of the larger of the two quantities divided by the difference between them. Writing X_1 as the larger and X_2 as the smaller of the quantities we thus have, if information is gained at a constant rate :

$$\text{Discrimination time} = K \log\left(\frac{X_1}{X_1 - X_2}\right) \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The similarity of this to eqn. (9) becomes clear when it is realized that movement time could have been described by the same equation with X_1 and X_2 representing respectively the distances to the far and near edges of the target.

Henmon's data fitted in terms of eqn. (10) are shown in Fig. 9. The linearity is striking except for a deviation by the left-most points for subjects S and W, and for subject H for discrimination between lines. This is perhaps due to some kind of lower limit to discrimination time in these cases analogous to that shown in Fitts's results for movement time. The slopes of the regressions for lines and tones for subject H are almost identical : a fact which suggests that the rate may be set by a central process common to discrimination in different sensory modes.

Birren and Botwinick's data fitted by eqn. (10) are shown in Fig. 10. The points understandably show greater deviation as they are based on fewer readings but are reasonably linear down to 10 to 15 per cent. From 15 to 50 per cent no further fall of time occurs : the limit to discrimination time hinted at in Henmon's results is here clearly in evidence. The two regression lines were fitted by least squares to the points down to 10 per cent. Their convergence to almost exactly the same point on the zero information line (at about 0·275 sec) is interesting when taken in connection with the fact that the difference between this point and the younger subjects' mean time for discrimination between 15 and 50 per cent is closely similar to the time commonly found for a two-choice reaction by people of their age. No data exists on choice reaction times for people as old as Birren and Botwinick's older subjects, but it is reasonable to assume that it would be about the value of the same difference in their case. If so, it looks as if discrimination of which line was longer and choice of response proceeded simultaneously, and that the recorded reaction time was either the choice time or the discrimination time, whichever was longer. Crossman (1955) found signs of the same relationship between discrimination and choice times in one of his experiments. The additional 0·275 sec was probably due to a time delay either in the apparatus or in the execution of the response.

One point should be noted about the finest discrimination, which is shown in Fig. 10 as 1·32 per cent. This was listed by the original authors as 1·0 per cent. Initial attempts by the present writer to fit the data all failed because this point consistently fell out of line. On a visit to Dr. Birren's laboratory, the writer was generously permitted to measure the lines used in the experiment and one of the two pairs nominally 1 per cent different was found to differ by slightly more than this. The fact that when the correct mean percentage difference was plotted the times were almost exactly those expected is a striking demonstration of the precision that is required in these experiments—a precision far greater than has been usual in psychological experiments in the past.

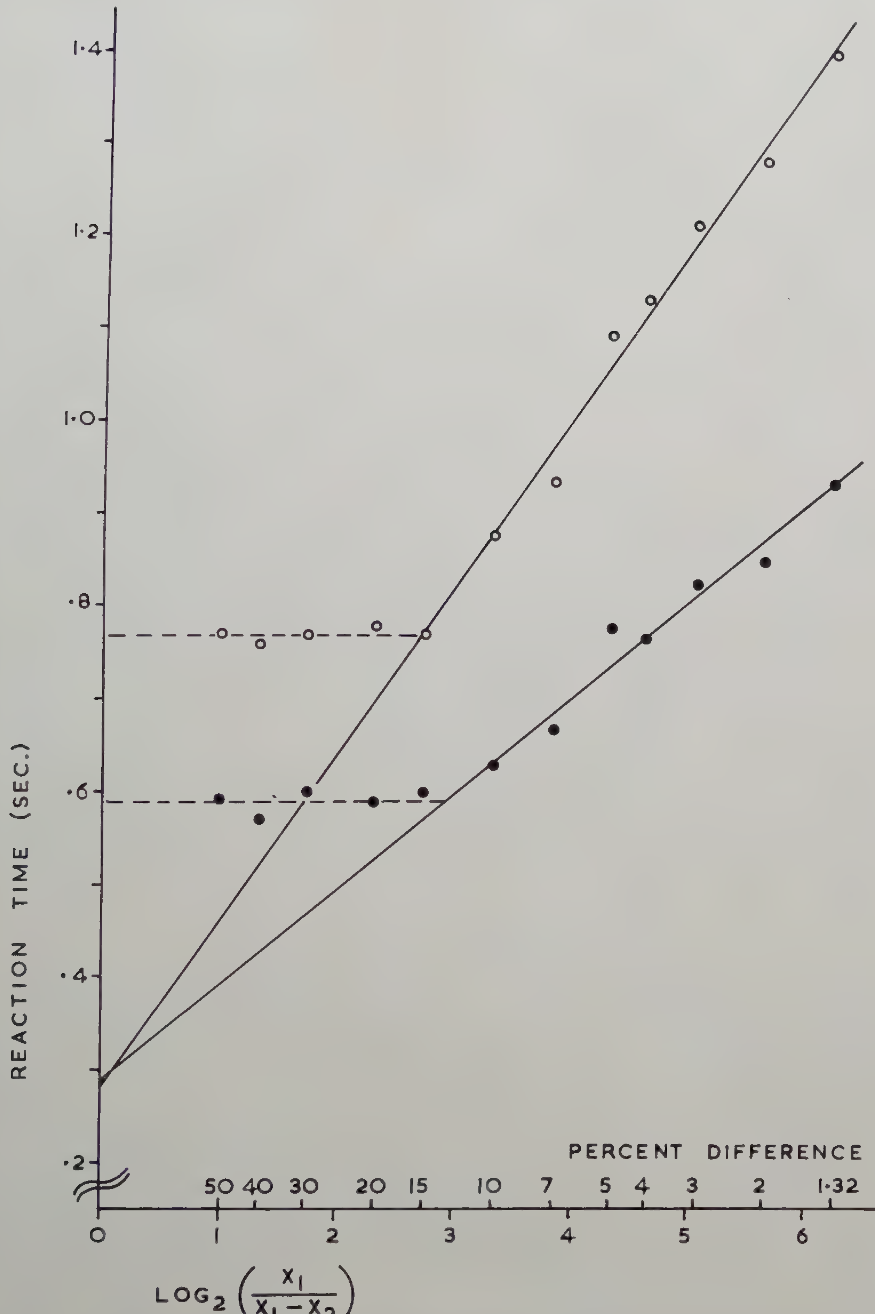

Figure 10. Birren and Botwinick's data plotted in terms of eqn. (10). Open circles, subjects aged 61–91. Filled circles, subjects aged 19–36. Each point is the mean of the medians of 43 older or 30 younger subjects. The medians were each based on at least four readings.

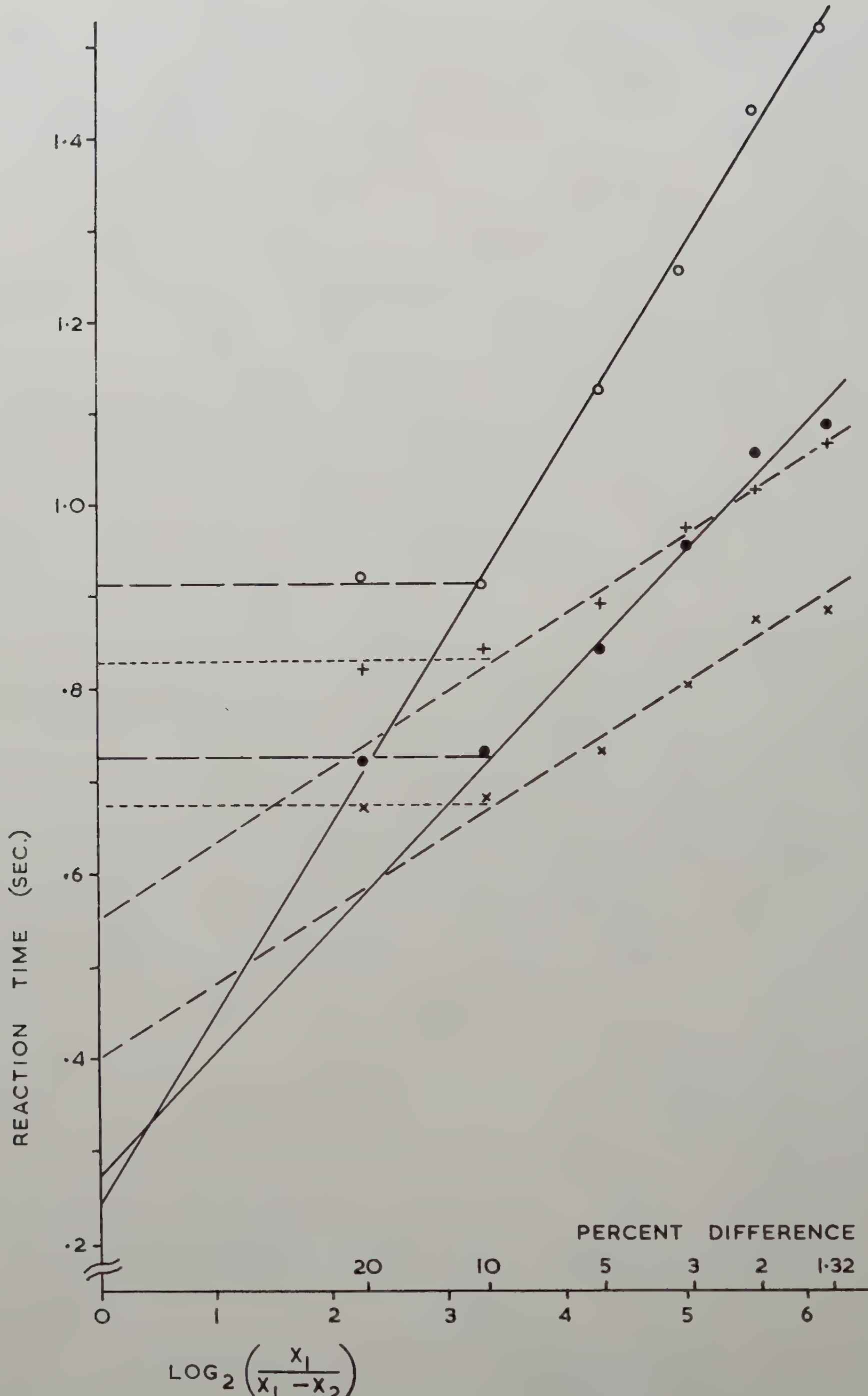

Figure 11. Data of Botwinick *et al.* plotted in terms of eqn. (10).
Results are indicated as follows :

Age range	Exposure time (sec)	Symbol
65–79	2·00	○
	0·15	+
18–35	2·00	●
	0·15	×

Each point is the mean of the medians of 34 older or 26 younger subjects. The medians were each based on eight readings.

The overlapping of discrimination and choice suggested by Birren and Botwinick's results appears to be of variable extent. The pattern of times found by Botwinick *et al.* and shown in Fig. 11 is closely similar to that of Fig. 10 for exposures of 2 sec but very different for exposures of 0·15 sec. In the latter case the slopes are lower and the regression lines cut the zero information line at different points, both above the meeting point of the regression lines for the results with the 2 sec exposure time. The fact that the regression lines of the two age groups are roughly parallel is consistent with the findings of Crossman and Szafran (1956) for different age groups performing a card-sorting task similar to Crossman's (1953). The general shortening of reaction times with the brief exposure is in line with results reported by Piéron (1920). It can be regarded as a way of optimizing performance since, when the signal has been removed, delay in responding cannot gain further data and will indeed tend to lose it owing to the decay or disruption of immediate memory.

Crossman found that eqn. (10) did not give a good fit to his own data, substantially underestimating the time required for finer discriminations. He, therefore, examined two further approaches.

(ii) *A statistical model.* If the nerve impulses representing the incoming signals are subject to greater or lesser supplementation from random ' noise ' impulses, it is reasonable to conceive of discrimination as a process of making a statistical test of the difference between two distributions (Tanner and Swets 1954; Gregory and Cane 1955; Crossman and Szafran 1956). Now the standard error of such a difference in normal statistical formulae varies as the square root of the number of readings. If we assume the subject to be taking ' readings ' at a steady rate, we may assume also that the achievement of discrimination to any given level of accuracy will vary as the square root of the time taken. More conveniently, we can postulate that the time taken will be a linear function of the square of the average of the reciprocals of the two Weber fractions $(X_1-X_2)/X_1$ and $(X_1-X_2)/X_2$ involved in comparing X_1 and X_2. Thus we can postulate that :

$$\text{Discrimination time} = K\left(\frac{1}{2}\,\frac{X_1+X_2}{X_1-X_2}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

One difficulty with this formula is that it postulates neural noise as operating upon the *ratios* instead of, as usually conceived, upon absolute magnitudes. This difficulty is not, however, serious as, if the concept of neural noise has any general validity, it must be assumed to enter to some extent at all stages of sensory-motor function from sense organs to action, and its operation to affect the ratios would mean no more than that it was entering at a relatively high level.

Crossman found this formula gave discrepancies from observation opposite to those of the information theory model, substantially *over*estimating the time for difficult discriminations.

(iii) *Crossman's ' confusion function '.* Crossman, in proposing a third formulation says, " we can consider S_1 and S_2 as points in a space of one dimension located at distances X_1, X_2 from the origin. The ease of distinguishing between them might then be expected to depend on the ' distance ' between them . . . . In order to make the distance dependent on ratio rather than on absolute difference we can take logarithms and measure in the space of Log x."

He goes on to suggest that the time taken for discrimination should be a linear function of the reciprocal of the differences between these logarithms so that :

$$\text{Discrimination time} = K\frac{1}{\log X_1 - \log X_2}. \qquad (12)$$

Now the expression $1/(\log X_1 - \log X_2)$ is approximately equal to $(X_1+X_2)/(X_1-X_2)$ multiplied by a constant. A proof of this is given in the Appendix. We can thus substitute for eqn. (12) the approximately equivalent, and rather more convenient equation

$$\text{Discrimination time} = K\left(\frac{(\frac{1}{2}X_1+X_2)}{X_1-X_2}\right) \qquad (13)$$

This does not differ appreciably from eqn. (12) except when X_1 and X_2 are very different. In these cases it corresponds closely to the revision of Crossman's formula proposed by Hammerton (1959).

Since

$$\frac{(X_1+X_2)}{X_1-X_2} = 2\frac{X_1}{X_1-X_2} - 1 = 2\frac{X_2}{X_1-X_2} + 1, \qquad (14)$$

discrimination time, if it is a linear function of $\frac{1}{2}(X_1+X_2)/(X_1-X_2)$, will also be a linear function of $X_1/(X_1-X_2)$ and of $X_2/(X_1-X_2)$. We may, in this connection note that substitution of $\log\{\frac{1}{2}(X_1+X_2)/(X_1-X_2)\}$ or $\log\{X_2/(X_1-X_2)\}$ for $\log\{X_1/(X_1-X_2)\}$ makes very little difference to the linearity of the results in Figs. 9, 10 and 11.

Some of Crossman's data are shown, fitted by eqn. (13) in Fig. 12. The linearity is quite good considering that the points are based on relatively small amounts of data.

Equation (13) does not give a good fit to the data of Figs. 9, 10 and 11, but it can be linked to the formula proposed by Piéron to describe the reduction of reaction time which occurs with increase of signal strength. Piéron (1920, 1936) reviewing his own work and earlier researches pointed out that simple reaction time can be thought of as made up of two parts : a portion which remains the same whatever the signal strength, and presumably has to do mainly with the translation from perception to action ; and a portion which varies with the strength of the signal. It seems reasonable to regard the latter portion as a discrimination time.

Piéron (1920) proposed the formula

$$\text{Reaction time} = \frac{a}{i+b} + c \qquad (15)$$

where i is the signal strength measured in physical units, and a, b and c are constants, the last representing the portion of the total reaction time which remains the same whatever the signal strength. Piéron found that eqn. (15) gave a reasonable fit to many sets of data with $b=0$. An example (Piéron 1920, p. 92) is given in Fig. 13. In these cases, if equation (15) with $b=0$ gives a linear fit to the data, the following will do so also :

$$\text{Reaction time} = K\frac{\frac{1}{2}(I+I')}{I \sim I'} + C \qquad (16)$$

where I is any ambient level of stimulation present before the signal arrives,

and I' is the level when the signal is present. This equation is clearly the same as eqn. (13) except for the constant C.

Piéron's eqn. (15) does not give a very good fit to much of the data, certainly not without an appreciable value of b, but the data cannot be pressed very far. We have already seen that extreme accuracy is needed and many of the experiments have been somewhat deficient in this respect. Pending further research,

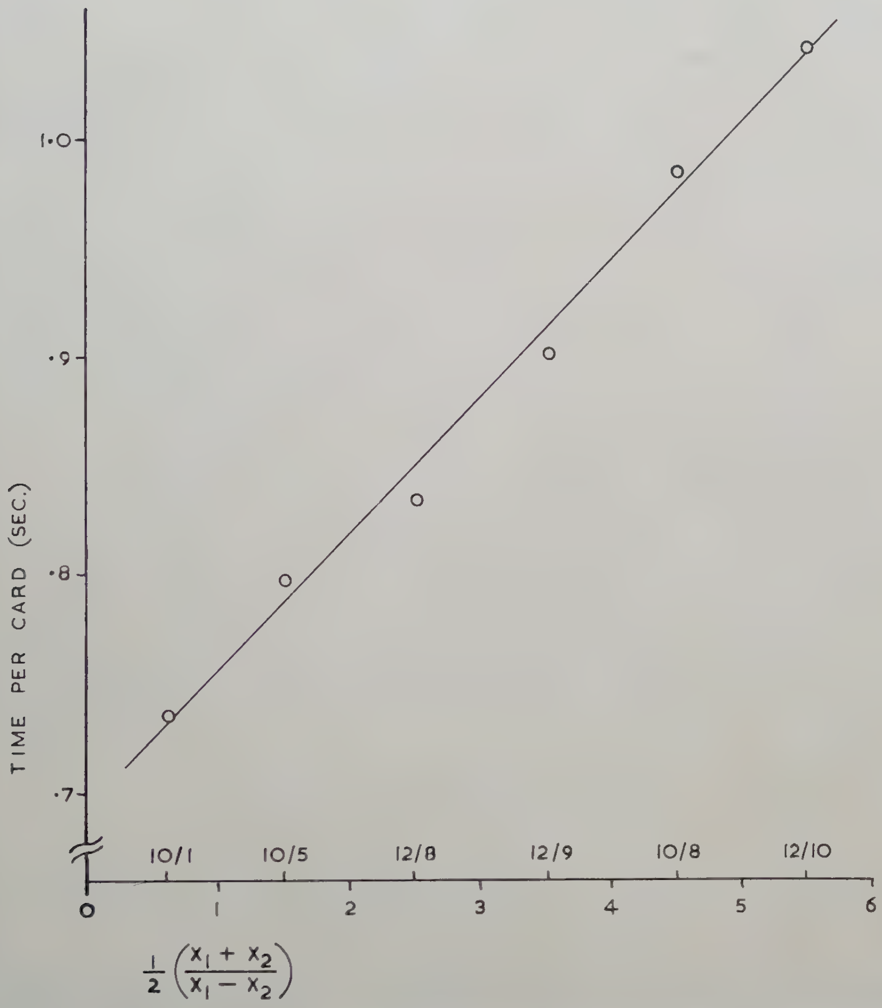

Figure 12. Crossman's data for sorting cards into two categories plotted in terms of eqn. (13). Each point is the mean time per card for four subjects each sorting 80–120 cards.

eqn. (16) can explain the results obtained by Steinman (1944) who found reaction times to a reduction of ambient stimulus level to be faster than those to an increase. For any given difference between I and I', the ratio $\frac{1}{2}(I+I')/(I \sim I')$ will obviously be larger and predict a longer reaction time if I' is greater than I than if it is smaller.

2.5.1. *Attempts to find an inclusive formula*

We may well ask whether there was any respect in which Crossman's experiments differed from those of Henmon, Birren and Botwinick and Botwinick *et al.* which could explain the difference of formula needed to fit them. The only consistent difference the present writer could detect was in the range over which discrimination was required. In Crossman's experiments the differences between X_1 and X_2 ranged from 90 to about 17 per cent, whereas in Henmon's they ranged from about 23 to 1·5 per cent, and in those of Birren and Botwinick and Botwinick *et al.* the linear regression in terms of eqn. (10) ranged from 10 or 15 to 1·32 per cent.

An extensive search was made of other possible formulations in the hope of finding one which would tie the two sets of results together. Two approaches appeared promising although neither was felt to be worth pursuing very far without new experiments.

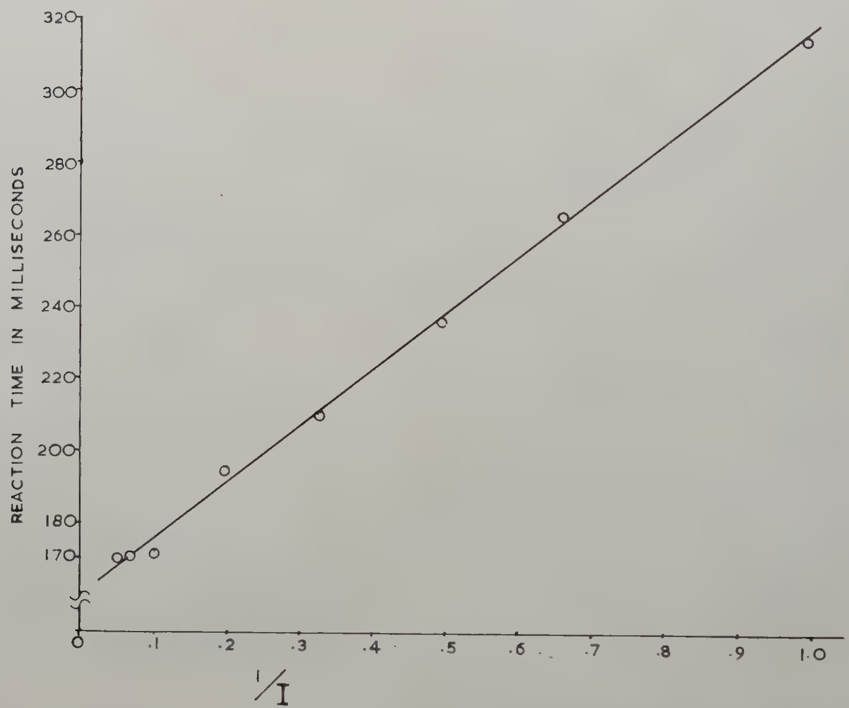

Figure 13. Data from an experiment by Piéron on the relationship between reaction time and the intensity of signal. The signal was a brief flash of light to a dark-adapted eye. Each point is the mean of 10–20 readings by one subject.

(i) *Allowance for error.* In all the sets of data examined errors tended to rise as the discrimination became finer. Adjustment of $(X_1 - X_2)$ in eqns. (10) and (13) to allow for these was tried in two ways:

(*a*) By the method proposed by Crossman (1957) for adjusting W in eqns. (8) and (9)—a method which assumes that the errors are distributed normally.

(*b*) By simply multiplying $(X_1 - X_2)$ by the proportion of responses made correctly—a method which assumes the errors to fall into a rectilinear distribution.

(*c*) A third attempt to adjust errors was made with eqn. (10) only and consisted of deducting from log $\{X_1/(X_1 - X_2)\}$ a measure of the information lost due to errors, i.e. $p_e \log (1/p_e) + p_c \log (1/p_c)$ where p_e = the proportion of errors and p_c = the proportion of correct responses. This method assumes, in the absence of precise evidence, that 'greater' was mistaken for 'smaller' as frequently as 'smaller' for 'greater'. It also assumes that the loss of information occurs in the process of choice rather than in discrimination. The choice might, perhaps, be deficient owing to the discrimination process in the perceptual mechanism not giving a clear enough signal to the translation mechanism.

It was thought inappropriate to adjust Crossman's results since the subject was required to correct any mistakes he noticed and errors were therefore unlikely to save him time. Adjustment of the other results to allow for errors was felt to be justified in that reaction time was measured to the first response, whether right or wrong. No adjustment could be made to Birren and Botwinick's results as no data were available about errors, but fits were tried to the results of Henmon and of Botwinick *et al.* The fits in terms of eqn. (10) adjusted by any one of the three methods were in all cases worse than if no correction had been made. Somewhat more promising results were obtained if plots were made in terms of eqn. (13) corrected for errors by methods (*a*) or (*b*), but the linearity was on the whole less good than in Figs. 9 and 11.

(ii) *A further possible formulation.* Of the attempts to fit the data with an entirely new formula the most promising used a combination of the two models that Crossman rejected. The rationale of this approach was that we can, perhaps, conceive of the subject, when making a discrimination, both as having to extract information in the sense of selecting one from a set of possible magnitudes, and also as having to resolve confusion due to random neural impulses which might add to and interfere with the signal. We could on this view write

$$\text{Discrimination time} = K\left[\log\left(\frac{X_1}{X_1 - X_2}\right)\right]^2. \quad \ldots \quad (17)$$

The fits obtained with this equation corrected for errors in appropriate cases were quite good to the results of Henmon, Botwinick *et al.* and Crossman. Those of Botwinick *et al.* are shown, corrected by method (*c*) in Fig. 14, and those of Crossman uncorrected for errors in Fig. 15.

The main difficulty with this formulation is that it appears somewhat far-fetched.

2.5.2. *Limits of discrimination*

What sets the ultimate limits to discrimination? It would seem arguable that, given sufficient time, there should be no limit to the fineness with which accurate discrimination can be made. Clearly this is not so, and the question arises of why not? Three possibilities appear worthy of consideration :

(*a*) There may be threshold factors due to neural noise. This has been held by several writers, notably Tanner and Swets (1954), to be the essential

determiner of absolute thresholds. Gregory and Cane (1955) have suggested that its effects on the differential threshold can be described by adding a small constant to the Weber fraction, i.e.

$$\frac{\delta I}{I+r}=K \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

where I is the physical intensity of a stimulus, δI is the least noticeable increase of stimulus, r is the constant representing the neural ' noise ' in terms of the stimulus, and K is also a constant. Such neural ' noise ' makes the expression $X_1/(X_1-X_2)$ become $(X_1+r)/(X_1+r)-(X_2+r)=(X_1+r)/(X_1-X_2)$, i.e. it adds only to the numerator of the expression.

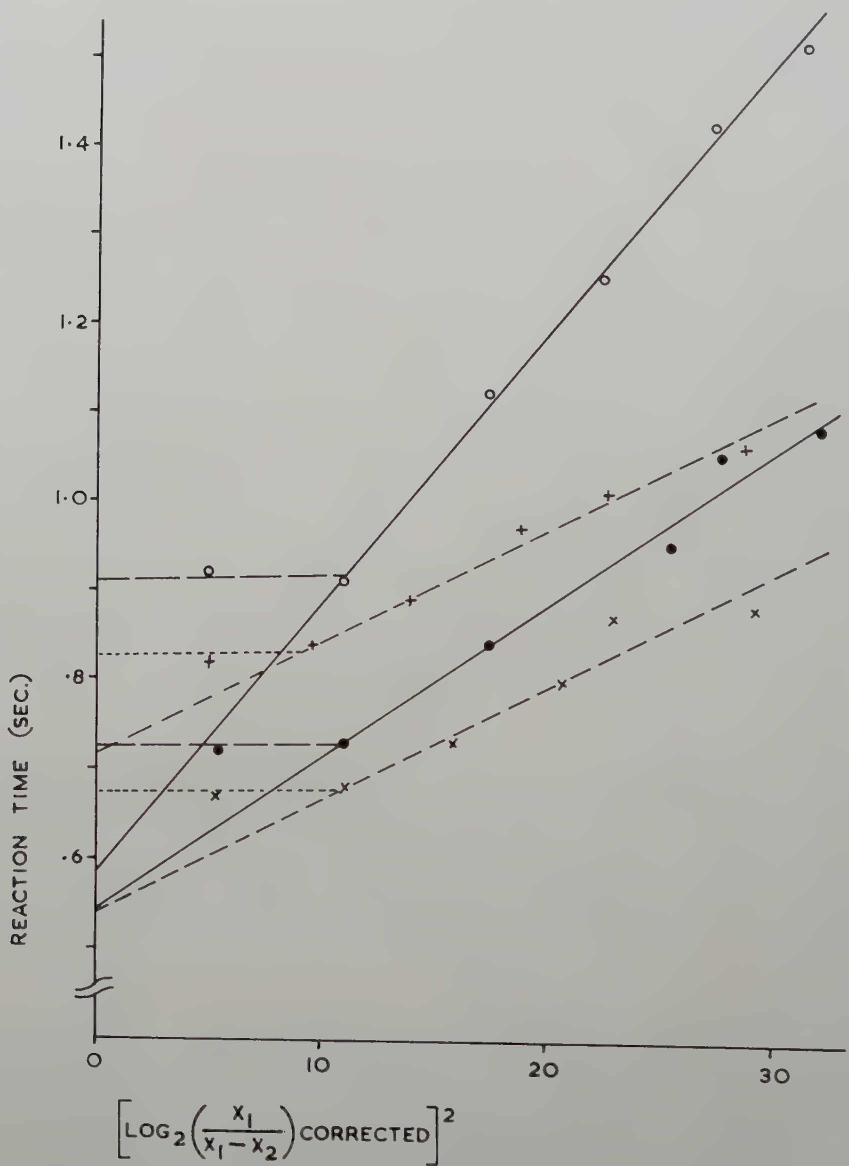

Figure 14. Data of Fig. 11 plotted in terms of eqn. (17) corrected for errors.

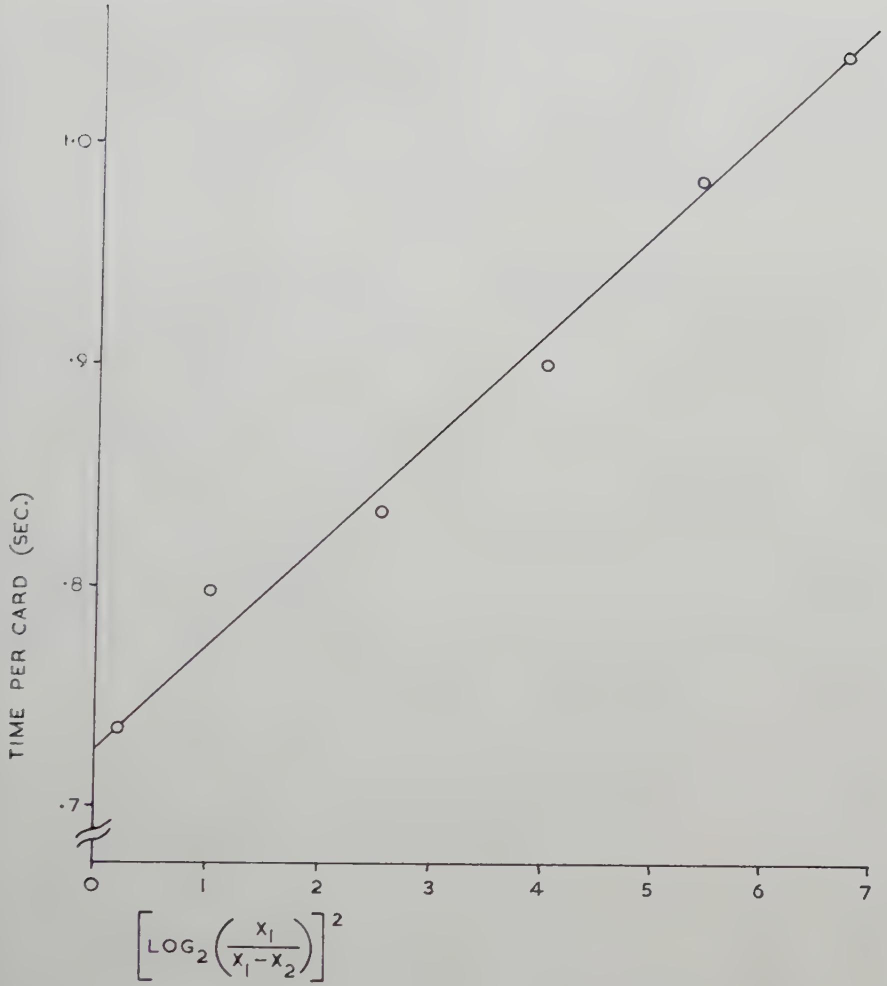

Figure 15. Data of Fig. 12 plotted in terms of eqn. (17)

Table 3. Times for discriminations of constant ratio but different absolute magnitude. Data from Henmon (1906) and Crossman (1955)

Henmon : discrimination between lines

	Lengths of lines (mm)	24/20	18/15	12/10	6/5
Subject H	Times (millisec)	303·85	307·73	309·71	314·98
	Errors (in 320 readings)	6	4	3	9
Subject S	Times (millisec)	323·00	323·50	328·50	337·40
	Errors (in 320 readings)	16	26	16	29

Crossman : Sorting packs of cards according to numbers of spots

Mean times per card for four subjects each sorting a pack of 40 cards three times at each ratio

Numbers of spots :	10/5	8/4	6/3	4/2	2/1
Times (millisec)	864	873	891	898	893

The data of Henmon and of Crossman in Table 3 for discriminations of constant ratio but different absolute magnitudes enable an assessment of these neural ‘ noise ’ effects to be made. In both sets of data the times tend to rise as the absolute magnitudes of the quantities being compared become smaller.

A rough calculation indicates that an r of about 0·2 mm for each of Henmon's subjects and about 0·25 of a spot in Crossman's experiment would account for the differences of times observed. The effects of these additions to the ratios in Henmon's and Crossman's experiments are obviously small. We do not know how far threshold effects were important in the other experiments but we may expect on the basis of Henmon's and Crossman's results that they were small also.

It is clear from eqn. (18) that the relative effects of neural noise become greater as the absolute magnitudes of the signals decrease. This view is supported by the results of Steinman (1944) who found that reaction time to any given change of stimulus intensity was greater when the absolute magnitude of the stimulus was small, than when it was larger. On the other hand eqn. (18) indicates that fine discriminations will not be much affected (Crossman and Szafran 1956). Neural ' noise ' of this kind does *not*, therefore, seem to be an adequate explanation of the limits of discriminability.

(*b*) As discrimination becomes finer the time required rises steeply and it may be that the subject is for some reason unable or unwilling to spend more than a given amount of time. Crossman (1955) suggested that, in his experiments, such a limit might lie at a discrimination time of about 0·5 sec. Such a limit might perhaps arise from expectations about ' normal ' time required, or pressure for speed exerted by the conditions under which a task was being done leading the subject to accept errors rather than spend more time.

(*c*) A third possibility arises from the fact that the accumulation of data over time in order to effect fine discrimination implies some form of short-term storage in which this data can be held until enough has been amassed. Such storage is almost certain to be ' leaky ' with the rate of loss rising with the amount of data in store. The longer the time for which data have to be held the more will, therefore, be lost, and eventually a point may be reached at which losses balance gains and the expenditure of even infinite further time will not result in any improvement of discrimination (for a discussion and tentative example of this type of model see Welford 1958, pp. 221–2). Such a model would require a shape of curve for relating time to very fine degrees of discrimination slightly different from any of those discussed here. The difference, would, however, be small.

2.5.3. *The need for further work*

It is clear that the data upon discrimination do not entirely agree and that, although eqns. (10), (13) and (17) can provide useful approximations for practical use, further research is required. The present discussion does, however, point clearly to certain requirements for this research. As a minimum we need

(i) full enough data for functional relationships between time and discrimination to be plotted over a very wide range,

(ii) a record of errors as well as of correct responses, and

(iii) readings in which the ratio of the quantities to be compared is kept constant and their absolute magnitudes are varied, as well as readings in which absolute magnitudes are kept approximately constant and ratios are varied.

In addition, we need a much fuller and more sensitive appraisal than has usually been made in the past of *what the subject is attempting to do*, of exactly what he is discriminating from what, of how he is setting about his task and how hard he is trying. Such observations are inevitably in part subjective and need to be checked against objective data, but they often provide an understanding of behaviour without which mathematical descriptions can be seriously misleading.

§ 3. Some Wider Implications of the Present Approach

Although the work outlined here has consisted of laboratory experiments, it would seem to have considerable implications for other activities, including some of practical importance which have hitherto eluded rigorous study. There is still a great deal to be done before these implications can be fully realized, but at this stage it seems fair to make six points :

1. Although the studies deal in the main with very brief performances such as pressing a key in response to a signal, the original stimulus for them was the attempt to analyse continuous performance and, in particular during the war and the immediate post-war years, the continuous following of a moving target. This interest led to a number of laboratory studies of tracking, and recently significant treatments of these in information-theory terms have been made by Griew (1958 a) and Crossman (1960). The conclusion that emerges from these studies and from consideration of the work as a whole is that action takes place in a matrix of *decisions* about signals and about actions rather than of *movements*, and that the limiting factors in the speed of an operation are the times taken by the central brain processes required for these decisions.

2. If this approach is to constitute a genuinely quantitative treatment of human behaviour, the rates of information-gain must be shown to transcend particular circumstances. Not enough is yet known for this to be done confidently, but a number of uniformities are emerging which suggest that the time when this can be done need not be long delayed. Some additional confidence may be gained from the point which has emerged from a number of studies (e.g. Singleton 1953; Welford 1959) that there is much greater uniformity between the reaction times of different individuals than between their times to make relatively unguided movements : in other words, the central mechanisms appear to function with less individual variation than do the peripheral.

3. In a general way the concepts developed over the years since the war provide a valuable new approach to job-analysis. The practical technique of this was pioneered by Crossman (1956 b) who has found it capable of dealing with a range of jobs not amenable to analysis by normal work-study methods. An example is *process* work such as in chemical and oil refining plants, where the operator may make many decisions but take few actions. This kind of work has sometimes been undervalued by previous methods of analysis and ratings leaning heavily on the study of movements, but it is now becoming possible to bring it into perspective. In some cases the demands of such work can be roughly calculated in advance. In other cases this is not feasible, but the extent to which an operator is ‘ loaded ’ by the decisions he has to make can still be estimated by adding subsidiary tasks of known information content

to his main task until signs appear of impaired performance. Data on the point at which breakdown begins make it possible, in principle, to calculate the demands of the main task alone. Exploratory examples of such studies at the laboratory level have been described by Mowbray (1952, 1953), Bahrick *et al.* (1954), Poulton (1958) and Griew (1959).

4. The results of the studies which have been outlined here have a number of implications for the design of work and machinery, especially for process plants and other forms of automation, but also for more conventional industrial equipment. The limited amount of decision and control of movement that can be achieved in any given time, means that the frequency and also the spacing of signals for action must be arranged so as to ensure that the operator is at no time overloaded. Paced work on conveyor lines is shown to be, in principle, inefficient (Conrad 1960) : variability of the time taken by successive cycles of an operation makes it inevitable that any fixed time per cycle which is long enough to include the longest of the operator's cycle times will leave him time to spare with the shorter, while if the fixed time is less than the longest a cycle will sometimes be missed or an error made.

5. The treatment of performance in the terms we have outlined, although it has, so far, been worked out only for simple sensory-motor tasks with implications mainly for shop-floor operations, seems likely also apply in principle to higher mental processes and thus to the work of administrators, executive staff and white-collar workers generally. The opening up of this line of study has still to be done, but it seems capable of providing a much more defined theoretical approach than has hitherto been available to certain social problems in industry, and indeed to some more general social problems.

6. Lastly, from the point of view of further theoretical research, it is becoming clear that although mathematical approaches, such as the application of information theory to psychology, are of very great importance, the formulations must not be made blindly. They are valid only in so far as they describe and codify the detailed behaviour of the subject, and are, in the long run, found to do the same for the physiological processes underlying it. The whole field seems, indeed, to be ready for a programme of research in which the contributions of experimental psychology and of physiology are closely integrated in the study of what has here been termed the *micro-behaviour* of decision processes.

My very sincere thanks are due to Dr. W. E. Hick for detailed criticism of two drafts of this paper, for a great many valuable ideas and for constant help and encouragement in discussion. Thanks are also due to Dr. J. Shakeshaft for the Appendix and to Dr. J. E. Birren, Dr. J. Botwinick, Dr. E. R. F. W. Crossman and Dr. J. A. Leonard for valuable discussions and for having brought to my attention a number of points that would otherwise have been missed. The paper is based upon a talk given to The Ergonomics Research Society on 7 April 1959.

Au cours des dernières années, on a reconnu davantage l'importance des activités perceptives et d'organisation centrale dans les performances sensori-motrices et on a fait des progrès vers un traitement authentiquement quantitatif. Cet article esquisse le développement historique du travail accompli dans ce domaine et tente une réévaluation autour de cinq idées principales :

(a) Il semble que les mécanismes centraux fonctionnent comme une voie unique qui ne peut

exploiter qu'un signal ou groupe de signaux à la fois, de sorte que des signaux se succédant rapidement doivent " faire la queue " avant d'y être utilisés.

(b) Les temps de réaction de choix sont commentés à la lumière de la théorie du gain d'information à un taux constant par le sujet. On discute aussi des modèles conceptuels du comportement de détail du sujet lorsqu'il effectue des choix.

(c) Des modèles concernant la rapidité et la précision des mouvements et qui sont issus de la théorie de l'information sont cités et discutés.

(d) Plusieurs formules sont examinées, tentant de relier le temps pris pour discriminer des quantités de différentes grandeurs avec la finesse de leurs différences.

(e) Un certain nombre d'extensions des travaux passés en revue sont suggérées. La plus importante sans doute est constituée par de nouvelles approches des tâches " mentales " et de surveillance, qui jusqu'à maintenant n'ont pas pu être soumises aux méthodes normales d'étude du travail.

On est amené à reconnaître, en conclusion, la nécessité de recherches à la fois psychologiques et physiologiques qui rendtraient possible le dépassement des formules mathématiques descriptives pour atteindre l'étude du micro-comportement de détail et de mécanismes neuromusculaires.

In den letzten Jahren wurde mehr und mehr erkannt, wie wichtig Aktivitäten der Wahrnehmung und der zentralen Organisation bei senso-motorischen Leistungen sind. In Richtung einer wirklich quantitativen Behandlung dieser Worgänge wurden Fortschritte gemacht. Diese Arbeit zeichnet die histroische Entwicklung auf diesem Gebiet und versucht eine Würdigung unter fünf Haupttiteln:

(a) Im zentralen Mechanismus tritt ein " Einzelkanal " auf, der sich nur mit einem Signal oder einer Gruppe von Signalen auf einmal beschäftigen kann. In rascher Folge eintretende Signale haben daher unter Umständen " Schlange zu stehen ", bevor sie verarbeitet werden.

(b) Wahl-Reaktionszeiten werden im Zusammenhang mit der Theorie diskutiert, dass eine Person Informationen mit konstanter Geschwindigkeit gewinnt (im Sinne der Terminologie der Informationstheorie). Passende Modelle des Einzel-Verhaltens einer Person, die eine Wahl trifft, werden gleichfalls diskutiert.

(c) Informations-Theorie-Modelle werden in Beziehung zur Geschwindigkeit und Genauigkeit von Bewegungen skizziert und diskutiert.

(d) Verschiedene Formeln werden geprüft, die versuchen, die Zeit für die Unterscheidung von Mengen verschiedener Grösse zur Feinheit ihres Unterschiedes in Beziehung zu setzen.

(e) Eine Reihe weiterer Beziehungen dieser Arbeiten werden herausgestellt und besprochen. Darunter sind vielleicht am wichtigsten neue Wege zum Studium der " geistigen " und überwachenden Aufgaben, die bisher mit den normalen Methoden der Arbeitsstudien nicht zu behandeln waren. Es wird geschlossen, dass gemeinsame psychologische und physiologische Forschung notwendig ist, die über descriptive mathematische Formeln hinaus das Studium des Mikro-Einzel-Verhaltens und neuro-muskulärer Mechanismen ermöglichen würde.

Appendix

Proof of the close similarity, except for a constant factor, between eqns. (12) *and* (13).

Putting T_D for discrimination time, eqn. (12) may be written :

$$T_D = K \frac{1}{\log_e\left(\dfrac{X_1}{X_2}\right)}$$

or

$$\log_e\left(\frac{X_1}{X_2}\right) = \frac{K}{T_D}. \qquad \text{(A 1)}$$

Now taking

$$T_D = K' \frac{X_1 + X_2}{X_1 - X_2}$$

$$= K' \frac{\dfrac{X_1}{X_2} + 1}{\dfrac{X_1}{X_2} - 1}$$

and rearranging
$$\frac{X_1}{X_2}=\frac{\frac{T_D}{K'}+1}{\frac{T_D}{K'}-1}$$

it can be shown then that
$$\log_e\left(\frac{X_1}{X_2}\right)=2\left\{\frac{K'}{T_D}+\frac{1}{3}\left(\frac{K'}{T_D}\right)+\frac{1}{5}\left(\frac{K'}{T_D}\right)^5+\;\ldots\;\to\infty\right\}.$$

Now if $X_1=X_2+\delta_X$ where $\delta_X \ll x_2$, then $(X_1/X_2)\to 1$ and $\log_e(X_1/X_2)\to 0$.
$\therefore\quad T_D\to\infty$ and $K'/T\ll 1$.

In this case $(K'/T_D)^3$, $(K'/T_D)^5$ and higher terms are negligible compared with (K'/T) and we are left with
$$\log_e\left(\frac{X_1}{X_2}\right)\simeq\frac{2K'}{T_D}$$
which is the same as eqn. (A 1) except for a constant factor.

References

Annett, J., Golby, C. W., and Kay, H., 1958, The measurement of elements in an assembly task—the information output of the human motor system. *Quart. J. exp. Psychol.*, **10**, 1–11.

Bahrick, H. P., Noble, M., and Fitts, P. M., 1954, Extra-task performance as a measure of learning a primary task. *J. exp. Psychol.*, **48**, 298–302.

Birren, J. E., and Botwinick, J., 1955, Speed of response as a function of perceptual difficulty and age. *J. Gerontol.*, **10**, 433–6.

Botwinick, J., Brinley, J. F., and Robbin, J. S., 1958, The interaction effects of perceptual difficulty and stimulus exposure time on age differences in speed and accuracy of response. *Gerontologia*, **2**, 1–10.

Bricker, P. D., 1955, Information measurement and reaction time. In *Information Theory in Psychology* (ed. H. Quastler) (Glencoe, Illinois : The Free Press), pp. 350–9.

Broadbent, D. E., 1952, Speaking and listening simultaneously. *J. exp. Psychol.*, **43**, 267–73 ; 1958, *Perception and Communication* (London : Pergamon Press).

Brown, I. D., 1960, Many messages from few sources. *Ergonomics*, **3**, 159–168.

Chapanis, A., Garner, W. R., and Morgan, C. T., 1949, *Applied Experimental Psychology* (New York : John Wiley & Sons).

Christie, L. S., and Luce, R. D., 1956, Decision structure and time relations in simple choice behaviour. *Bull. math. Biophys.*, **18**, 89–111.

Conrad, R., 1960, Letter sorting machines—paced, ‘lagged’ or unpaced? *Ergonomics*, **3**, 149–157.

Craik, K. J. W., 1943, *The Nature of Explanation* (Cambridge University Press) ; 1947, Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, **38**, 56–61; 1948, Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.*, **38**, 142–8.

Crossman, E. R. F. W., 1953, Entropy and choice time : the effect of frequency unbalance on choice response. *Quart. J. exp. Psychol.*, **5**, 41–51 ; 1955, The measurement of discriminability. *Ibid.*, **7**, 176–95; 1956 a, The information capacity of the human operator in symbolic and non-symbolic control processes. In Ministry of Supply Publication WR/D2/56, *Information Theory and the Human Operator ;* 1956 b, Perception study—a complement to motion study. *The Manager*, **24** (2), 141–5 ; 1957, The speed and accuracy of simple hand movements. In *The Nature and Acquisition of Industrial Skills* by E. R. F. W. Crossman and W. D. Seymour, Report to M. R. C. and D. S. I. R. Joint Committee on Individual Efficiency in Industry; 1960, The information-capacity of the human motor-system in pursuit tracking. *Quart. J. exp. Psychol.*, **12**, 1–16.

Crossman, E. R. F. W., and Szafran, J., 1956, Changes with age in the speed of information intake and discrimination. *Experientia Supplementum*, **4**, 128–35.

Davis, R., 1956, The limits of the ‘psychological refractory period’. *Quart. J. exp. Psychol.*, **8**, 24–38 ; 1959, The role of “attention” in the psychological refractory period. *Ibid.*, **11**, 211–20.

Denton, G., 1953, Times spent handling and transporting shoes. Brit. Boot, Shoe and Allied Trades Res. Ass. Rep. No. TM 1123.

FISHER, R. A., and YATES, F., 1957, *Statistical Tables for Biological, Agricultural and Medical Research* (5th ed.) (London : Oliver and Boyd).

FITTS, P. M., 1947, *Psychological Research in Equipment Design* (Washington : U.S. Govt. Printing Office), 1954, The information capacity of the human motor system in controlling the amplitude of movement. *J. exp. Psychol.*, **47**, 381–91.

FRAISSE, P., 1957, La période réfractoire psychologique. *Année Psychol.*, **57**, 315–28.

GIBBS, C. B., 1952, Car turning signals and delays in responding to visual information. *M.R.C. Appl. Psychol. Res. Unit Rep.* No. 124/50.

GREGORY, R. L., and CANE, V., 1955, A statistical information theory of visual thresholds. *Nature, Lond.*, **176**, 1272.

GRIEW, S., 1958 a, Age changes and information loss in performance of a pursuit tracking task involving interrupted preview. *J. exp. Psychol.*, **55**, 486–9 ; 1958 b, Uncertainty as a determinant of performance in relation to age. *Gerontologia*, **2**, 284–9 ; 1959, Set to respond and the effect of interrupting signals upon tracking performance. *J. exp. Psychol.*, **57**, 333–7.

HAMMERTON, M., 1959, A mathematical model for perception and a theoretical confusion function. *Nature, Lond.*, **184**, 1668–9.

HENMON, V. A. C., 1906, The time of perception as a measure of differences in sensation. *Arch. Philos. Psychol. Sci. Meth.* No. 8.

HICK, W. E., 1948, The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36–51 ; 1952 a, On the rate of gain of information. *Ibid.*, **4**, 11–26 ; 1952 b, Why the human operator ? *Trans. Soc. instrum. Technol.*, **4**, 67–77.

HOUSEHOLDER, A. S., and LANDAHL, H. D., 1945, *Mathematical Biophysics of the Central Nervous System* (Bloomington, Indiana : Principia Press).

HYMAN, R., 1953, Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.

KLEMMER, E. T., 1957, Simple reaction time as a function of time uncertainty. *J. exp. Psychol.*, **54**, 195–200.

LEMMON, V. W., 1927, The relation of reaction time to measures of intelligence, memory and learning. *Arch. Psychol., N.Y.*, **15**, No. 92.

LEONARD, J. A., 1959, Tactual choice reactions I. *Quart. J. exp. Psychol.*, **11**, 76–83.

MARILL, T., 1957, The psychological refractory phase. *Brit. J. Psychol.*, **48**, 93–7.

MERKEL, J., 1885, Die zeitlichen verhältnisse der willensthätigkeit. *Philos. Stud.*, **2**, 73–127.

DE MONTPELLIER, G., 1935, *Les Altérations Morphologiques des Mouvements Rapides* (Louvain : Institut Supérieur de Philosophie).

MORIN, R. E., and GRANT, D. A., 1955, Learning and performance on a key-pressing task as a function of the degree of spatial stimulus-response correspondence. *J. exp. Psychol.* **49**, 39–47.

MOWBRAY, G. H., 1952, Simultaneous vision and audition : the detection of elements missing from overlearnt sequences. *J. exp. Psychol.*, **44**, 292–300 ; 1953, Simultaneous vision and audition ; the comprehension of prose passages with varying levels of difficulty. *Ibid.*, **46**, 365–72.

MOWBRAY, G. H., and RHOADES, M. V., 1959, On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, **11**, 16–23.

PIÉRON, H., 1920, Nouvelles recherches sur l'analyse du temps de latence sensorielle et sur la loi qui relie ce temps à l'intensité de l'excitation. *Année Psychol.*, **22**, 58–142 ; 1936, Recherches sur la latence de perception des accroissements de luminosité. *Ibid.*, **37**, 1–16.

POULTON, E. C., 1958, Measuring the order of difficulty of visual-motor tasks. *Ergonomics*, **1**, 234–9.

RABBITT, P. M. A., 1959, Effects of independent variations in stimulus and response probability. *Nature, Lond.*, **183**, 1212.

RAPOPORT, A., 1959, A study of disjunctive reaction times. *Behavioural Sci.*, **4**, 299–315.

SCHOUTEN, J. F., VREDENBREGT, J., and ANDRIESSEN, J. J., 1960, On the laws governing the times needed for pin-mounting as a function of diameter and tolerance. Paper to Ergonomics Res. Soc. 29 March 1960. Abstract : *Ergonomics*, **3**, 000–000.

SHANNON, C. E., and WEAVER, W., 1949, *The Mathematical Theory of Communication* (Urbana, Illinois : University of Illinois Press).

SIMON, J. R., and SIMON, BETTY P., 1959, Duration of movements in a dial setting task as a function of the precision of manipulation. *J. appl. Psychol.*, **43**, 389–94.

SINGLETON, W. T., 1953, Deterioration of performance on a short-term perceptual-motor task. In *Symposium on Fatigue* (ed. W. F. Floyd and A. T. Welford) (London : H. K. Lewis for the Ergonomics Research Society).

STEINMAN, ALBERTA R., 1944, Reaction time to change compared with other psychophysical methods. *Arch. Psychol., N.Y.*, **41**, No. 292.

Tanner, W. P., and Swets, J. P., 1954, A decision-making theory of visual detection. *Psychol. Rev.*, **61**, 401–9.

Venables, P. H., 1958, Stimulus complexity as a determinant of the reaction time of schizophrenics. *Canad. J. Psychol.*, **12**, 187–90.

Vince, Margaret A., 1948 a, The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, **38**, 149–57 ; 1948 b, Corrective movements in a pursuit task. *Quart. J. exp. Psychol.*, **1**, 85–103 ; 1949, Rapid response sequences and the psychological refractory period. *Brit. J. Psychol*, **40**, 23–40 ; 1950, Some exceptions to the psychological refractory period in unskilled manual responses. *M.R.C. Appl. Psychol. Res. Unit Rep.* No. 124/50.

Wehrkamp, R. A., and Smith, K. U., 1952, Dimensional analysis of motion : II. Travel distance effects. *J. appl. Psychol.*, **36**, 201–6.

Welford, A. T., 1952, The ‘ psychological refractory period ’ and the timing of high-speed performance–a review and a theory. *Brit. J. Psychol.*, **43**, 2–19 ; 1958, *Ageing and Human Skill* (Oxford University Press for the Nuffield Foundation) ; 1959, Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, **11**, 193–210.

Wiener, N., 1948, *Cybernetics,* (New York : Technological Press of M.I.T. and John Wiley and Sons.)

Woodworth, R. S., 1938, *Experimental Psychology* (New York : Henry Holt & Co.).

Woodworth, R. S., and Schlosberg, H., 1954, *Experimental Psychology* (Revised ed.) (New York : Henry Holt & Co.).

A COMPARISON OF THREE DIAL SHAPES FOR CHECK-READING INSTRUMENT PANELS

By K. F. H. Murrell*

(Formerly of the Naval Motion Study Unit)

Three dial shapes for check reading were compared in an experiment in which three matched groups of 24 subjects each watched one panel for two periods of 90 minutes. Results show that, using either speed of response or number of stimuli missed as the criterion, a dial with a fixed pointer and a moving scale was inferior to a dial with a fixed scale and a moving pointer. There was a significant difference in response time between vertical linear and circular dials in favour of the former. But this result may be suspect because of the apparent unsatisfactory matching of the women subjects. There was no significant difference in the numbers of missed stimuli on these two dials. It is concluded that either circular or edgewise dials may be used in panels according to circumstances. The results of a supplementary experiment with two subjects lasting 15 days suggested that deviations on dials in the top left and bottom right of the panels are noticed more quickly than those in the other parts.

§ 1. Introduction

When a panel of dials has to be watched for long periods, it may be the task of the watchkeeper to notice when the value indicated on any one of them deviates from a fixed reading or range of readings, and then to take appropriate action. Thus the primary purpose of the panel should be for ease of check reading and secondly for ease of qualitative reading (rate and direction of change). Speed of quantitative reading (the actual value of an indication), is usually of little importance in this situation.

Most panels currently in use employ conventional circular dials arranged so that all zeros are in the ' 7 o'clock ' position ; their pointers when giving a ' normal ' reading are usually in random positions and no indication is given of the ' normal ' readings. The learning time on such a panel is, as a result, long and monitoring can be unnecessarily fatiguing. Warrick and Grether (1948) have shown that some improvement can be effected by turning the scales so that pointer alignment in the ' normal ' position is in the 9 o'clock position. Their test panel of 16 instruments was ' check read ' in 0·8 seconds as compared with 1·6 seconds for random alignment. When an ' up or down ' response was required the total time for the 9 o'clock panel was approximately 3·0 seconds. The comparable time for mixed alignment was 5·1 seconds approximately.

This method of aligning the pointers (which is mechanically not very expedient), coupled with a clear marking of the correct reading will produce a very marked improvement in ease and speed of reading but it will not get over two major disadvantages of the circular dial. The first is that a circular dial occupies a large area of the panel surface for the information it must display. The effect of this is to make some panels very wide and the resulting eye movements required in scanning very extensive, thus it follows that the scanning time for a wider panel is greater than that required for a narrower panel. This has been confirmed by White (1949).

* Now Research Fellow in Psychology in the University of Bristol.

Secondly, a circular dial is ‘ one-sided ’ when it is desired rapidly to comprehend the direction of change. In an experiment on dial shapes, Grether and Connell (1948) found that readings on the left side of a circular dial were slightly faster and much more accurate than readings on the right (0·6 seconds to 0·9 seconds : 6·25 per cent to 32·8 per cent errors) ; they suggested that this is because a movement upward of the pointer on the right side indicates a decrease in value which is incompatible with the ‘ expected ’ direction of movement.

It is therefore desirable to consider some alternative form of dial which may prove superior to the circular dial for check, qualitative or quantitative reading. Sleight (1948) showed that for quantitative reading five shapes could be ranked in the following order : (*a*) open-window (*b*) round (*c*) semi-circular (*d*) horizontal and (*e*) vertical. Experiments by Grether and Connell (1948) on single instruments suggested that moving pointer instruments are in general superior to moving scale instruments for check and qualitative readings ; that a circular moving pointer instrument is slightly superior to a linear moving pointer instrument for check reading, but that the order is reversed for qualitative readings. On the other hand Sleight (1949) showed that moving scales are superior to fixed scales and that small open window dials are preferable to a clear scale for ‘ speed and accuracy of reading ’. Connell (1950) using panels of four dials showed that for check reading a moving pointer, centered at 9 o’clock, is superior to a moving scale instrument or a direct reading counter.

It is difficult to decide from the findings of these workers which type of instrument may be substituted for the circular dial in a panel of multiple dials ; a decision which is not made any easier by the fact that the experiments on which these findings are based were in the main carried out on single dials using a tachistoscopic method. They may, therefore, bear very little relationship to the design of a panel of multiple dials for check reading in a watchkeeping situation where no action is required until something deviates from normal.

It was, therefore, decided to carry out an experiment in a watchkeeping situation with three dial shapes, circular with moving pointer, vertical linear with moving pointer, vertical linear with moving scale.

§ 2. The Experiment

2.1. *Apparatus*

Three simulated panels were prepared, each containing 18 dials arranged in three rows of 6. Panel A (Fig. 1) consisted of circular dials of 8·2 cm in diameter spaced 5 cm apart. Panel B consisted of drums 8·2 cm in diameter with a scale 1·6 cm wide graduated round them ; they were viewed through windows 2 cm × 5 cm spaced 8 cm apart. The pointers made straight lines across the face of the panel (Fig. 2). Panel C consisted of scales similar to those in Panel B curved to a radius of 4·1 cm and spaced 8 cm apart. The pointers moved in an arc close to the surface of the scale (Fig. 3). The panels were matt black ; the scales were black on white and all had the same gradation interval of 5 mm. On each a ‘ safe area ’ was marked in green, varying in

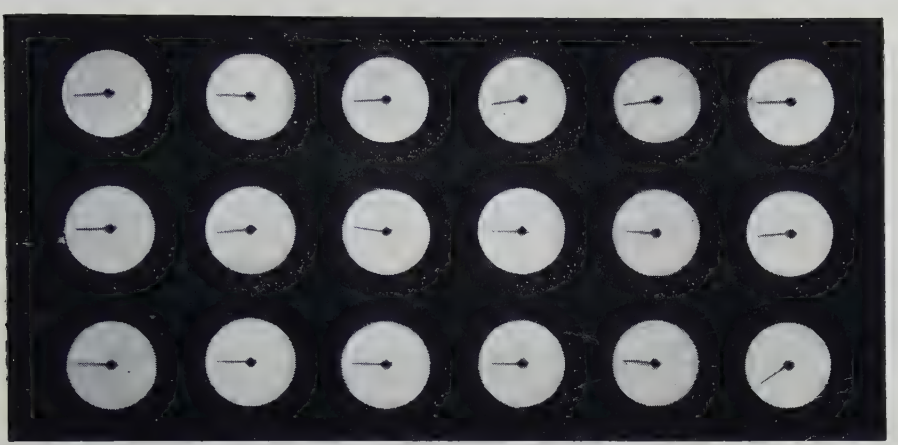

Figure 1.

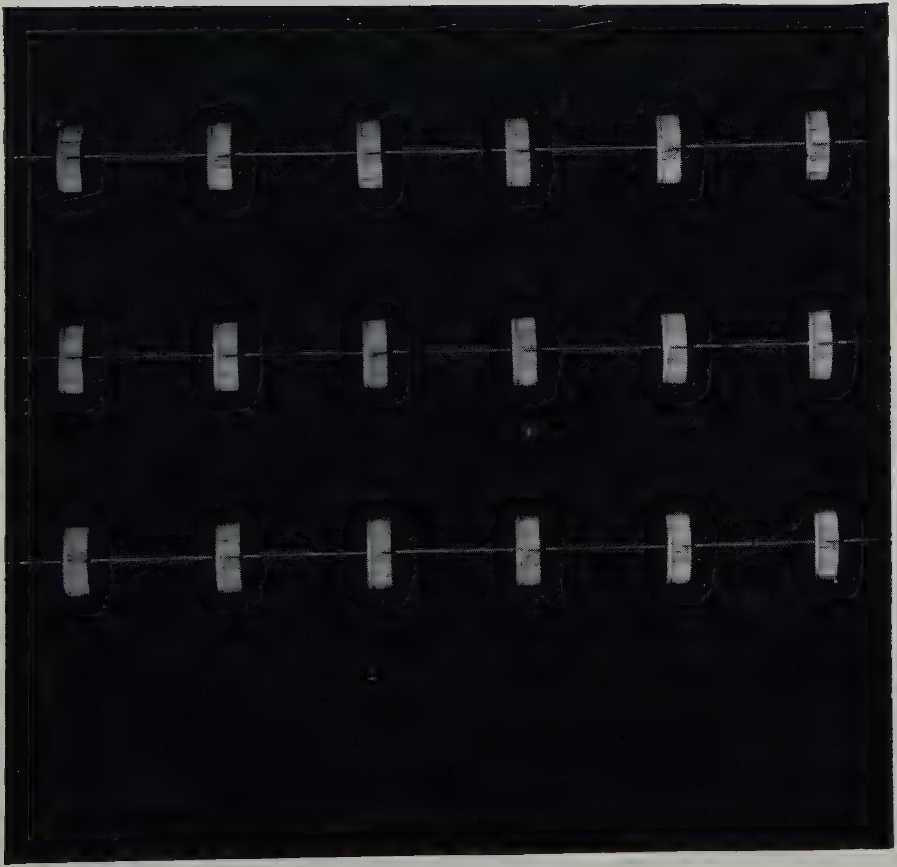

Figure 2.

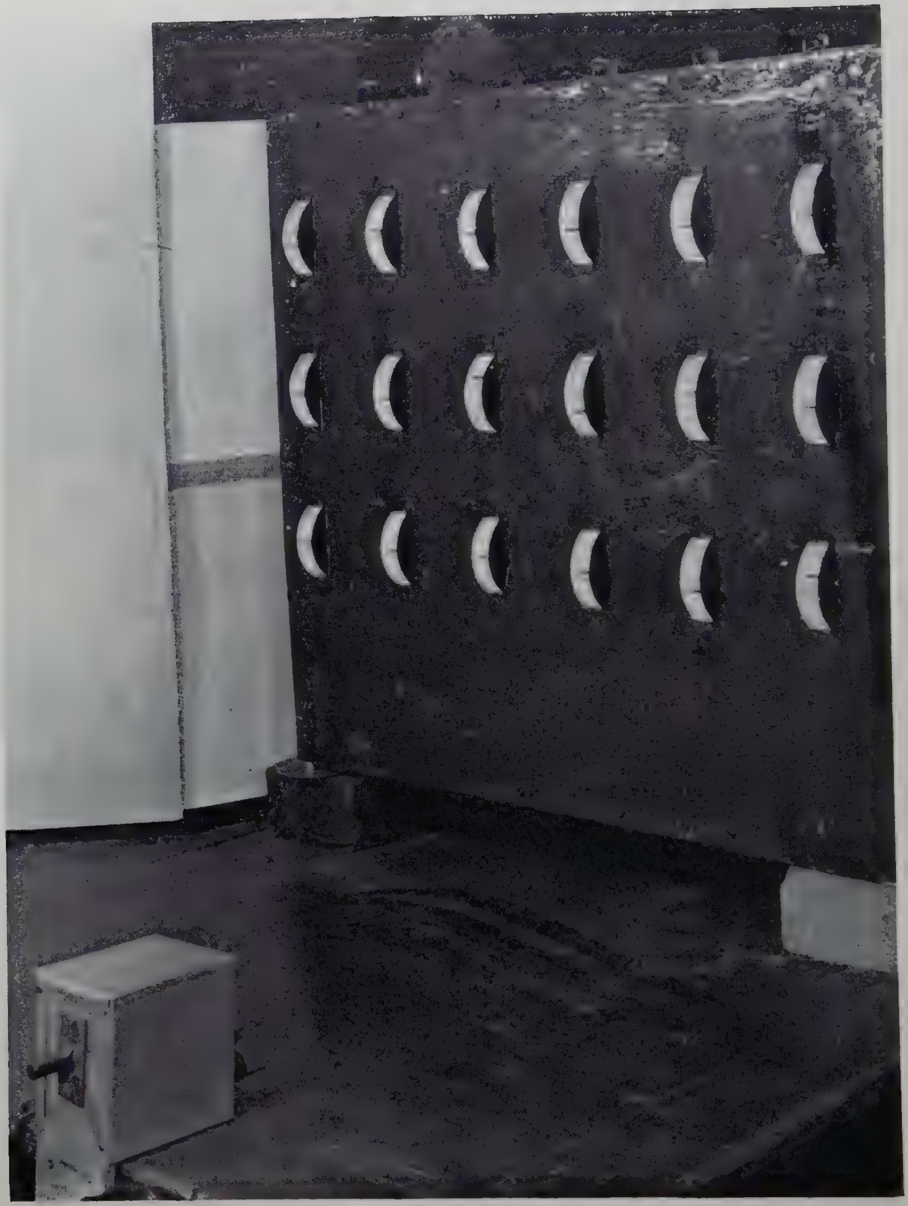

Figure 3.

width from 2–14 scale divisions. Each dial was exactly similar to those in the same position on the other two panels.

The pointers were driven at three speeds so that they moved continuously up and down inside their ‘ safe areas ’. The experimenter could cause any pointer to be driven outside its ‘ safe area ’ ; when this occurred, a wiper on the back of the panel closed a contact to give a record of both time and direction. The subject, on seeing that a pointer had deviated, moved a response key up or down according to the meaning of the indication. This was also recorded

and thus it was possible to measure the direction both of the stimulus and the response and the elapsed time between them.

The subject sat in a chair 45 cm high with his eyes 100 cm away from the panels. The centres of the panels were at the average eye-level. Two 2 ft fluorescent tubes were mounted vertically close to the subject's shoulders, giving even illumination of 12 foot candles at the surface of each panel.

The motors driving the panels provided sufficient background noise to mask the movements of the experimenters and to simulate conditions in a control room. The level measured at the position of the subject's head was 80–82 d.b. (referred to 0·0002 dynes/cm^2).

In order that the speed of movement of the pointers should have no undue influence, the contacts were calibrated with reference to a point 1·5 mm outside each safe area, a distance which was determined objectively to be the minimum distance at which the pointers could be seen clearly to be ' out '.

2.2. *The Subjects*

A heterogeneous group of 72 subjects were used, drawn from Scientific and Clerical grades of the Royal Naval Scientific Service and the Department of Medical Director General. Each subject was available for only a few hours, which ruled out the use of an experimental design involving many replications. To eliminate possible transfer effects, the 72 subjects were divided into three groups of 24, each of which was tested on a single panel. As far as possible the groups were matched on variables which were expected to be correlated with performance in the experiment (tests, age, sex and Civil Service grade), but naturally, it was impossible to match perfectly on all independent variables. The tests (N.I.I.P. G.T. 70/23—non-verbal intelligence ; SP Test 119—scale and dial reading ; SP Test 20—checking, speed and accuracy) were given by members of the Senior Psychologist's Department, Admiralty, who also did the matching.

2.3. *Procedure*

Each subject did two 1½-hour sessions, one in the forenoon and one in the afternoon of the same day. Fifteen stimuli (deviations from ' normal ') were given in each session ; the intervals between them varied between 2–10 minutes and were with one exception either short (2–4 minutes) or long (8–10 minutes). The dials on which the stimuli were given were selected to give 5 stimuli each in the inner, middle, and outer columns and 5 each with fast moving, medium, and slow moving pointers ; these, with direction of movement, were given in random order by ' drawing out of a hat '.

The subjects were given typewritten instructions varying slightly from panel to panel ; for example those for the circular instruments read :

> " You have before you a panel of circular instruments in which the scale is fixed and the pointer moves. The scale is graduated in a clockwise direction and the pointer moves clockwise to increase the value. On each scale is marked a green ' safe band ' and the object of the experiment is to record how quickly and accurately you notice when one of the indicators reads a value outside this ' safe band '.

You have a key in front of you which will move either up or down. If you notice that an instrument is reading high, i.e. if the pointer has moved outside the 'safe band' in a clockwise direction, you should *immediately* push the key up; similarly, if you notice that an instrument is reading low, you should push the key down. The pointers will be moving throughout the experiment and the high and low readings will appear at irregular intervals. You are asked to watch as carefully as possible and to give the response as quickly as possible."

A demonstration of the panels was then given and when the subjects signified that they were ready and understood what they had to do the run was started.

§ 3. Results

There are two criteria on which the relative merits of the three panels may be judged: the response time and the missed stimuli. The exact degree of importance of each must depend upon the particular situation under which a panel is to be used. The results will be dealt with separately.

There are a number of other effects which might have been tested which would be of interest but which have no important bearing on the primary object of the investigation; but owing to the heterogeneous nature of our experimental subjects and the lack of complete matching in our groups, any effects shown might be peculiar to our particular experiment. It may be said, however, that we have been unable to find any effect due to differences in dial position, i.e. by comparing vertical columns or horizontal lines, or any difference due to pointer speed, width of 'safe area' or extent of pointer deviation.

Table 1. Response time

Panel		Men			Women			All Subjects		
		a.m.	p.m.	Both sessions	a.m.	p.m.	Both sessions	a.m.	p.m.	Both sessions
A	$\bar{X}$	6·52	6·42	6·47	8·81	8·15	8·47	7·42	7·15	7·33
	S.D.	1·464	1·737	1·634	1·816	1·296	1·655	2·033	1·798	1·882
B	$\bar{X}$	7·42	8·98	8·17	7·77	9·38	8·63	7·56	9·16	8·36
	S.D.	1·845	1·506	1·768	2·626	2·492	2·744	2·287	2·127	2·223
C	$\bar{X}$	6·58	6·11	6·34	6·89	6·75	6·81	6·71	6·39	6·54
	S.D.	0·974	2·016	1·613	1·689	1·616	1·703	1·368	1·909	1·599
Significance $p \leqslant$										
A—B		NS	0·01	0·001	NS	NS	NS	NS	0·001	0·001
A—C		NS	NS	NS	0·05	0·05	0·01	NS	NS	0·05
B—C		NS	0·01	0·001	NS	0·02	0·02	0·05	0·001	0·001

3.1. *Response Times*

The mean response times and the standard deviations for men and women separately, and together, for both replications and for the whole experiment are given in Table 1, and are shown graphically in Fig. 4.

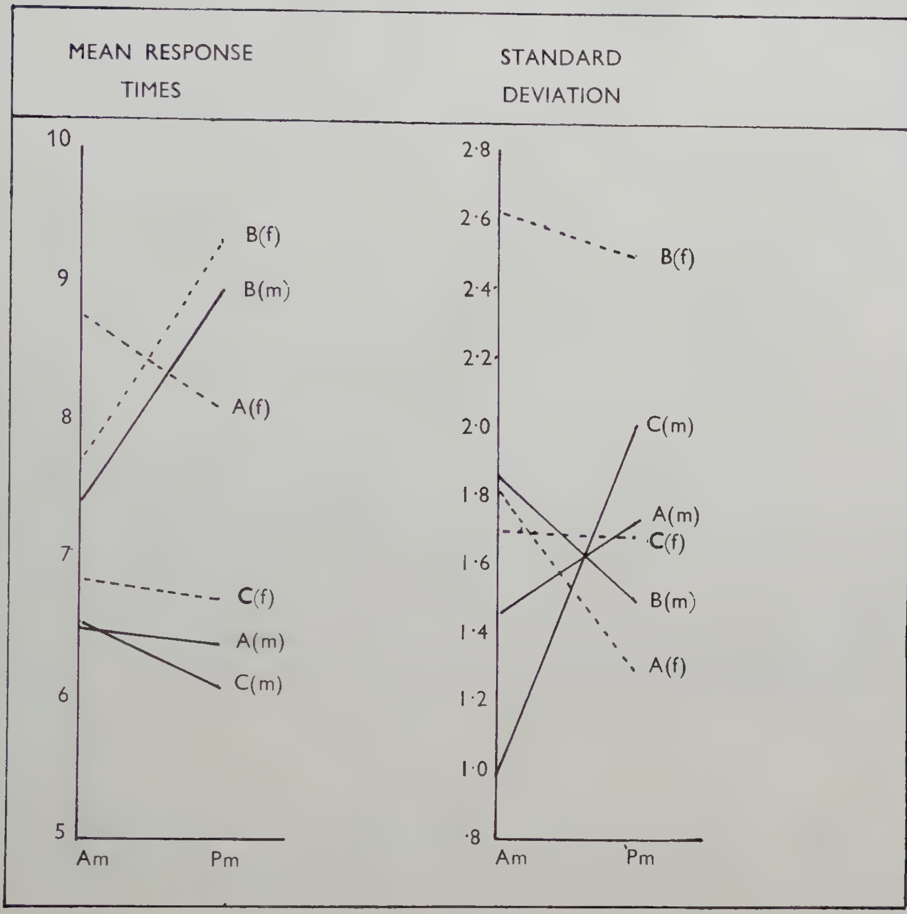

Figure 4.

The women were slower than the men, the greatest difference (2·0 sec) being on Panel A, whereas on the other two panels the differences between the sexes were almost the same at ·46 and ·47 sec. This difference on Panel A may be surprising since it would seem that circular dials with a moving pointer would have been a familiar shape even to the most naïve of subjects. The men did not seem to experience any difficulty with this panel, but both men and women clearly found Panel B unsatisfactory. Panel C seems to give uniformly the best results with both sexes and over the experiment as a whole : it is read significantly faster than the other two panels. But this result seems to be almost entirely due to the curious difficulty which the women had with Panel A. In any experiment using reputedly matched groups a result which depends on the performance of one group of the subjects may perhaps be suspect and due consideration must be given to the rather large difference between the men and women on Panel A.

There was a general slight but non-significant improvement in response times in the afternoon sessions except on Panel B on which both men and women were much slower, which emphasizes the difficulty which the subjects had in using this type of display.

There were large differences between the response times of the various subjects and these are reflected in the rather large standard deviations. Any slight practice effect between the two sessions might be expected to show a reduction in the S.D. and this is found in the women ; for the men, on the other hand, there was substantially greater scatter on both Panels A and C. It may be that the men were more likely to have eaten a larger lunch between the sessions, but if they did it did not cause them to go to sleep since there were less stimuli missed by the men on both these panels in the afternoon. Alternatively, they could have ‘ eased off ’ a little at the prospect of a second rather boring session and this may account for the slight reduction in the S.D. on Panel B which may have required the most effort.

3.2. *Missed Stimuli*

The number of missed stimuli (‘ misses ’) in the different sessions are shown in Table 2, together with the number of individuals who missed one or more stimuli. It will be seen, that except for the men, on Panel B less stimuli were missed by fewer individuals in the afternoon than in the forenoon. Taking both sexes together there were an equal number of ‘ misses ’ in the forenoon on Panels B and C, with Panel A somewhat less, but in the afternoon Panel C had improved so that it is now equal to Panel A. On the experiment as a whole there is no significant difference between Panels A and C but Panel B is significantly worse than Panel A at ·05 per cent level.

Table 2. Missed stimuli

		Males			Females			All Subjects		
		a.m.	p.m.	Both	a.m.	p.m.	Both	a.m.	p.m.	Both
A	Total	15	9	24	14	11	25	29	20	49
	$\overline{X}$	1·07	0·64	0·83	1·4	1·1	1·25	1·21	0·83	1·03
	N	8	5		5	5	10	13	10	
B	Total	18	19	37	19	8	27	37	27	64
	$\overline{X}$	1·29	1·36	1·33	1·9	0·8	1·35	1·54	1·13	1·33
	N	8	10		6	4		14	14	
C	Total	18	13	31	19	7	26	37	20	57
	$\overline{X}$	1·29	0·93	1·10	1·9	0·7	1·30	1·54	0·83	1·19
	N	8	7		7	5		15	12	

$\overline{X}$=Mean number of ‘ misses ’ for all subjects.
N=Number of subjects who missed one or more stimuli.

Whilst there is little evidence of any learning effect between sessions from the response times, there were substantially less ' misses ' (67) in the afternoon than there were in the forenoon (103), the difference being significant at $p < \cdot 01$. This would suggest that the subjects are becoming more accustomed to working on the panels, Panel C showing the greatest improvement. From this it might be argued that misses may perhaps be regarded as a more sensitive measure of the relative merit of a design than response time.

The relationship between speed of response and number of misses can be considered either in relation to panels and sessions or in relation to the individual subjects. If the panels/sessions/sexes are ranked by mean response times and by number of ' misses ' there is found to be very little relationship between the two rankings ($\tau = 0 \cdot 333$ which is quite insignificant). If the subjects are divided into the half returning the fastest times and the half returning the slowest times, the number of ' misses ' associated with each are given in Table 3. Taken as a whole the slower subjects missed many more stimuli than did the faster subjects ($\chi^2 = 18 \cdot 44 : p < \cdot 001$) but this result is more due to the men ($\chi^2 = 17 \cdot 39 : p < \cdot 001$) than to the women ($\chi^2 = 3 \cdot 28 : p < \cdot 07$). If the table is consulted it will be seen that it is the women on Panel A who are the cause of this difference (for Panels B and C only, $\chi^2 = 6\ 81 : p < \cdot 01$). From this it must perhaps be reluctantly concluded that the women were not adequately matched and that the performance of those on Panel A was inferior to that expected on the basis of the performance on the other two panels. This is one of the risks which has to be taken in an experiment using matched groups and although some of the results are suspect it does not mean that the experiment is valueless. In the first place, even with the women, the actual difference between Panels A and C is quite small and secondly, all three panels can still be compared on the basis of the men and Panels B and C on the basis of all subjects.

Table 3. Relationship between individual speed of response and stimuli missed

Subjects		Panel A	Panel B	Panel C	All Panels	
Fastest half	♂	10	7	9	26	57
	♀	14	8	9	31	
Slowest half	♂	14	30	22	66	113
	♀	11	19	17	47	

Table 2 also shows that nearly half the subjects monitored the panels without any ' misses ' at all, a result which would not have been predicted on the basis of laboratory research on vigilance. The duration of each deviation

was 25 sec and this may have influenced the result although the findings of Bakan (1955) and Saldanha (1956), suggest that this may not necessarily be so ; the frequency of the deviation (15 in 90 min) may also be a contributory cause, but again this does not differ greatly from stimulus frequency in other vigilance experiments. An alternative possibility is that our subjects, drawn from the Civil Service and working on a task which bore some resemblance to real life, were better motivated than the sailors or students usually used as subjects in vigilance experiments. Whatever the cause, the fact remains that nearly half the subjects showed no decrement in vigilance over $1\frac{1}{2}$ hours of the experiment and this emphasizes the difficulty of extrapolating with confidence from laboratory experiments to work in control rooms and elsewhere.

§ 4. Conclusion

Panel B clearly is inferior to the other panels on both criteria. The difference in response times on Panels A and C of 0·79 sec, although statistically significant is likely to be of relatively little practical importance and on the basis of the men alone there is no significant difference between the panels. Missed stimuli are, however, a matter of great importance ; using ‘ misses ’ as a criterion there is no difference between Panels A and C, but it is interesting to speculate whether the greater improvement of Panel C would have been maintained had the experiment continued longer, making it superior to Panel A.

We can, then, conclude that under the conditions of the experiment dials with fixed scales and moving pointers are superior to dials with fixed pointers and moving scales ; there is no practically important difference between edgewise dials and circular dials, although the response times on the former are significantly the faster. Thus, in constructing panels either shape can be used with confidence although preference might well be given to edgewise dials when a panel of minimum width is required. With either shape the chance of missing a deviation is altogether too great and some supplementary warning would seem to be imperative if the consequences of failure to notice a deviation are likely to be costly.

§ 5. A supplementary experiment

In addition to the analysis given above, which covers the main object of the experiment, the variation of performance with clock time on each of the three panels was also of interest. The mean reaction time for each of the six 15-minute periods was computed for each panel ; these are given in Fig. 5. The performance on Panels A and B followed a normally accepted course, but on Panel C the times for the second 45 minutes were appreciably faster than those for the first 45 minutes. This effect was found to have occurred equally in both sessions and was shown by approximately 75 per cent of the subjects.

It was decided, therefore, to investigate this further in a supplementary experiment with two subjects. In this experiment, a 90-minute session was taken once by each subject on 15 successive days (omitting week-ends). Subject NGS started one day ahead of NJB who had acted as subject in the main experiment. Their average reaction times for each day are given graphic-

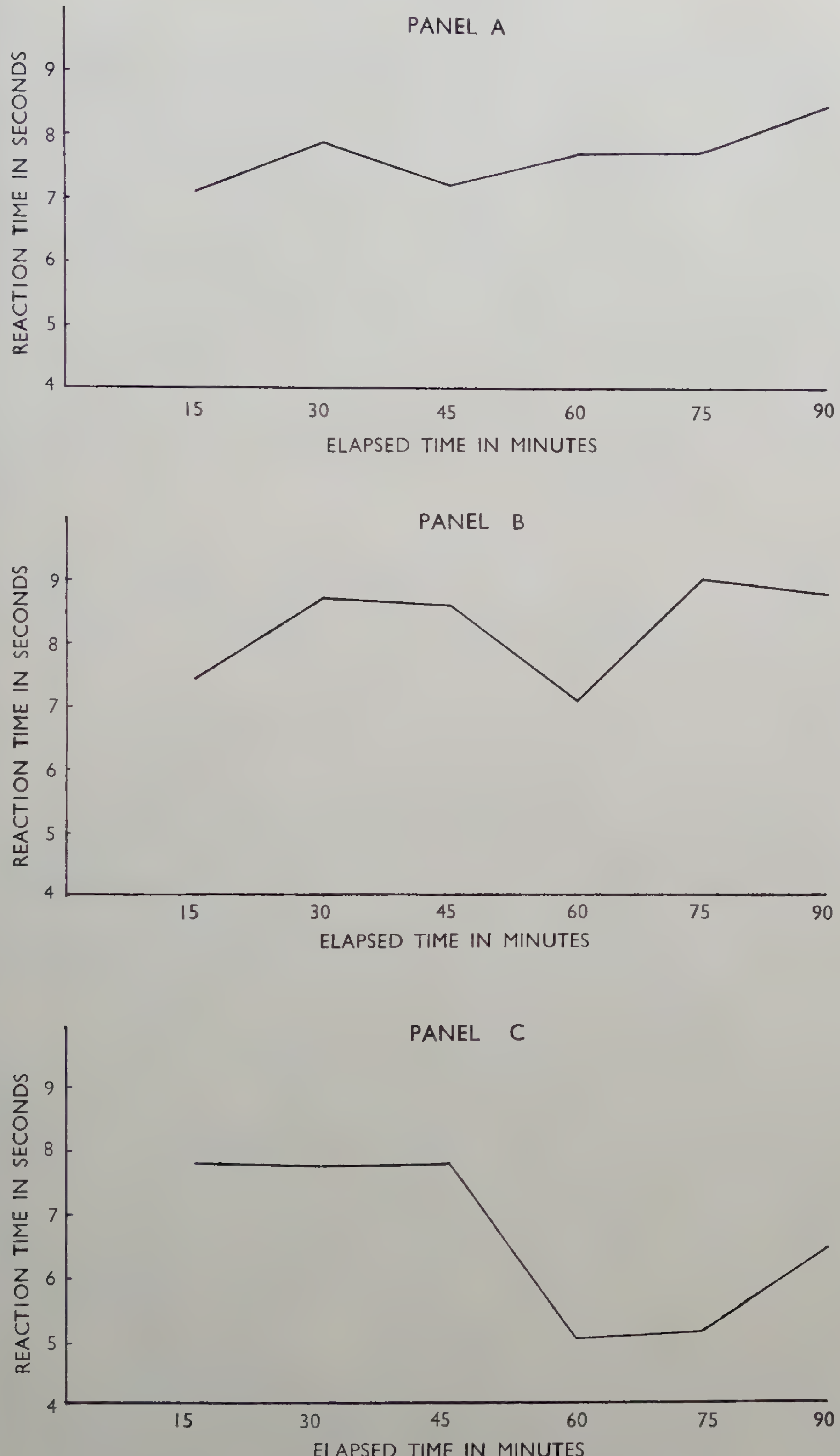

Figure 5. Average performance in each 15-minute period.

ally in Fig. 6, which shows that up to the eighth day there was an improvement in their performance, after which they appeared to level off. This improvement was not due to learning the sequence of the stimuli, because a different sequence was used during each session within the limitation that the group of dials used during the first 45 minutes of the main experiment (group A) were not mixed with the dials used in the second half of the main experiment (group B). From session to session, the order in which group A or group B was given was changed—sometimes group A was given first and sometimes group B. The average reaction times for each group are given in Table 4 ; the differences between the groups are given on the hypothesis that the reaction times for Group B are faster than the reaction times for group A, in whatever order the groups are given. On this hypothesis, the mean reaction times for group B are significantly faster than those for group A at the ·01 per cent level of confidence ($n = 29$, $t = 3{\cdot}715$).

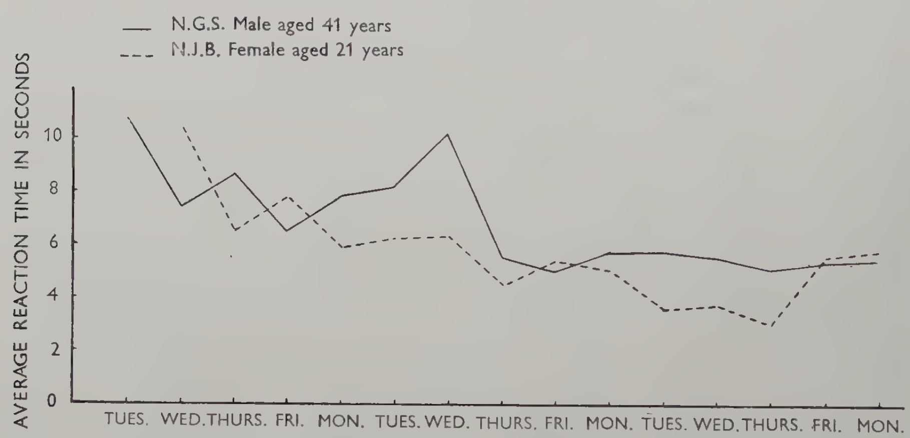

Figure 6. Average reaction times on successive days during the supplementary experiments.

Table 4. Mean reaction times in supplementary experiment

Day	Subject NJB					Subject NGS				
	1st 45 min		2nd 45 min		A—B	1st 45 min		2nd 45 min		A—B
	A	B	A	B		A	B	A	B	
1						12·9			9·0	3·9
2	11·4			9·4	2·0	10·4			5·3	5·1
3		5·4	7·4		2·0		10·7	6·9		−3·8
4		7·4	6·9		−0·5		6·5	6·4		−0·1
5	5·4			6·5	−1·1	11·0			5·1	5·9
6	6·7			5·5	1·2		7·6	8·4		0·8
7		5·3	7·3		2·0		9·1	11·3		2·2
8	4·7			4·4	0·3	7·0			4·5	2·5
9		4·2	6·6		2·4		3·6	6·4		2·8
10	6·1			4·0	2·1	5·8			5·9	−0·1
11		4·3	3·3		−1·0		5·7	6·1		0·4
12		3·8	3·9		−0·1		5·7	5·6		−0·1
13	4·4			1·8	2·6		4·0	6·2		2·2
14	6·7			4·4	2·3	5·8			4·8	1·0
15		4·8	6·6		1·8	4·9			5·6	−0·7

The panel was carefully examined to discover why this should be so ; the speed of the pointer movement, the width of the safe area, the distance by which the pointers deviated from normal, and the direction of movement of the pointers were all examined but none of these appeared to account for the effect found. It was found, however, that the dials which made up group B were those in the top-left and bottom-right hand corners of the panel ; and this positional effect appeared to be the only possible explanation. It is curious that the same effect was not found on either of the other panels.

An additional point which may be noted is the difference between the subjects. Over the last eight trials (when performance seems to have levelled off) the performance of NGS, aged 41, is significantly slower than that of NJB, aged 21, on the basis of the Mann-Whitney test ($U=12$: $p=\cdot 025$). In addition NJB showed the differences between the dial groups more markedly than did NGS. On the basis of the Wilcoxon Signed Ranks test NJB showed the difference with a probability of $p=\cdot 01$ whereas NGS showed a probability of $p=\cdot 05$. This may be an age effect.

The permission of the Lords Commissioners of the Admiralty to publish this paper is gratefully acknowledged. S. J. Murch, N. G. Suffield and Nita J. Morley all took an active part in conducting the experiment. The matching of the groups was undertaken by members of the Senior Psychologist's staff.

Trois formes de cadrans d'appareils indicateurs à surveiller dans un panneau de contrôle ont été comparées au cours d'une expérience dans laquelle trois groupes appariés de 24 sujets surveillèrent chacun un panneau pendant deux périodes de 90 minutes. Les résultats montrent que, utilisant comme critères soit la vitesse de réponse, soit le nombre d'omissions, un cadran avec index fixe et échelle mobile était inférieur à un cadran avec échelle fixe et index mobile. Il y avait une différence significative des temps de réponse entre les cadrans linéaires verticaux et les cadrans circulaires en faveur de ces derniers. Mais ce résultat est suspect du fait d'un appariement insuffisant des sujets de sexe féminin. Il n'y avait pas de différence significative quant au nombre d'omissions entre ces deux types de cadrans. En conclusion, les deux types de cadrans, circulaires ou linéaires, peuvent être montés dans des panneaux de contrôle, selon les circonstances. Les résultats d'une expérience complémentaire effectuée sur 2 sujets pendant 15 jours montre que les écarts survenant dans les cadrans de l'angle supérieur gauche et de l'angle inférieur droit du panneau sont repérés plus rapidement que ceux des autres régions du panneau.

Instrumenten-Tafeln mit 3 verschiedenen Ziffernblattformen für Kontroll-Ablesungen wurden in einem Experiment verglichen, in dem jede von 3 gleichartigen Gruppen von 24 Personen eine Tafel 2 Perioden von 90 min überwachte. Benutzt man entweder die Geschwindigkeit der Reaktion oder die Zahl der verpassten Reize als Kriterium, so zeigt sich, dass ein Ziffernblatt mit festem Zeiger und beweglicher Skala einem Ziffernblatt mit fester Skala und beweglichem Zeiger unterlegen ist. Es bestand eine signifikante Differenz in der Reaktionszeit zwischen vertikal—linearen und zirkulären Zifferblättern zugunsten der ersteren. Doch dieses Ergebnis ist unsicher wegen ungenügender Gleichartigkeit der weiblichen Versuchspersonen. Es bestand keine signifikante Differenz in der Zahl der verpassten Reize. Es wird geschlossen, dass auf Instrumenten-Tafeln je nach den Umständen zirkuläre oder lineare Ziffernblätter benutzt werden können. Ein zusätzliches Experiment mit 2 Personen von 15 Tagen Dauer machte es wahrscheinlich, dass Abweichungen auf Ziffernblättern oben links und unten rechts auf einer Tafel schneller bemerkt werden als an anderen Stellen.

References

Bakan, P., 1955, Discrimination decrement as a function of time in a prolonged vigil. *J. exp. Psychol.*, **50**, 387–390.

Connell, S., 1950, Some variables affecting instrument check reading. *USAF Technical Report No.* 6024, August 1950.

Grether, W. F. and Connell, S. C., 1948, Psychological factors in check reading of single instruments. *USAF Air Material Command Report* MC REXD 694. 17a, September 1948.

Saldanha, E. L., 1956, An investigation into the effects of prolonged and exacting visual work. *MRC App. Psychol. Unit Report No.* 243.

Sleight, R. B., 1948, The effect of instrument dial shape on legibility. *J. app. Psychol.*, **32,** 170–188.

White, W. J., Warwick, M. J., and Grether, W. F., 1953, Instrument reading III. Check reading of instrument groups. *J. app. Psychol.*, **37,** 302–307.

White, W. J., 1949, The effect of dial diameter on ocular movements and speed and accuracy of check reading of simulated engine instruments. *USAF Technical Report No.* 5826.

A NOTE ON PRINTING TO MAKE COMPREHENSION EASIER

By E. C. Poulton

Medical Research Council, Applied Psychology Research Unit, Cambridge

This paper collects together and evaluates the experimental evidence upon the relation of style of printing to the rate of comprehension of the reader. Difficulties in interpretation are mentioned, and a suggestion is made as to the kind of research which now most needs to be carried out.

Table 1 summarizes the recommendations on printing made by three authors (Luckiesh and Moss 1942 ; Burt, Cooper and Martin 1955, also Burt 1959 ; and Tinker and Paterson 1940, also Tinker 1960).

Table 1

	Recommended Optimal Printing		
	Burt	*Tinker and Paterson*	*Luckiesh and Moss*
Type Style	Old Style best, except possibly for scientists	Only American Typewriter and Cloister Black not recommended	Modern better than Old Style
Type form	—	Roman better than all Capitals or Italics	Italics insufficiently bold
Boldness	Only if eye defects	Unnecessary	Medium Bold recommended
Size	10 point	10 point	At least 12 point
Leading between lines	1 or 2 point	2 point	3 point or more
Length of line	3½–5½ inches	3 inches	2 inches or less
Colour of print and paper	—	Dark letters on light paper best. Contrast alone matters	Ditto ; non-glossy paper and ink recommended

Considering the relatively small range of conditions within which the optimal style of printing must necessarily lie, the table shows that there could hardly be greater disagreement between the different authors. The greatest disagreement is between Burt at one extreme and Luckiesh and Moss at the other ; the recommendations of Tinker and Paterson tend to be intermediate.

It is possible to see how the discrepancies could have arisen when the evidence upon which the recommendations in Table 1 are based is examined in detail. *Burt's* experiments with adults apparently did not produce results which were reliable statistically (1959, p. 10). In view of this he appears to have interpreted his results in the light of his own impressions. His recommendation on the boldness of type is the only one in Table 1 which is not apparently backed by any experiments of his own. In this case he refers to an experiment by Luckiesh and Moss, but refuses (perhaps rightly) to be influenced by the results (1959, p. 10 footnote 2).

Luckiesh and Moss who were lighting engineers, used rate of blinking as their principal criterion of readability. They did not attempt to measure comprehension, and, probably as a result of this, the rate of reading hardly varied with the difficulty. Tinker (1948 a) has shown that rate of blinking is not correlated with rate of comprehension. From Luckiesh and Moss's experimental results it would appear that rate of blinking is a measure of *visibility* rather than of comprehensibility. Their recommendations thus apply more to conditions of reading in which visibility is critical, such as poor eyesight, reading at long distances or with very poor light, than to good conditions of reading in which it is the rate of comprehension which matters most.

Tinker and Paterson measured rate of reading while comprehension remained above a fixed level. Each comparison was carried out on a separate group of about 80 students, and the differences found were subjected to acceptable statistical tests. Recommendations are based upon statistically reliable differences in the rate of comprehension. They are therefore the only recommendations in Table 1 which are at all acceptable for ease of comprehension under ordinary conditions of reading. Even these recommendations must, however, be interpreted in the context of their rather restricted experimental conditions, and should be used only with caution when conditions of reading are different. Four limitations upon Tinker and Paterson's recommendations are apparent :

1. One limitation concerns the *population* from which the experimental subjects were drawn : they used American college students, mainly freshmen. As Burt points out (1959, p. 18), people may read most easily the style of printing to which they are accustomed. If American printers use rather different styles from those used in Britain, Tinker and Paterson's results may not be fully applicable in this country.

2. A second limitation depends upon Tinker and Paterson's practice when comparing two conditions on each experimental subject of always presenting the *standard condition*, to which all the experimental variables are to be related, *before* the variable condition. The size of the differences found along a dimension may thus depend upon the nature of the standard condition with which each variable in the dimension was compared. For example when a line of 10-point Scotch Roman $3\frac{1}{8}$ in. long without leading was presented first as a standard, they found that it was read 7·3 per cent faster than a similar line $2\frac{1}{4}$ in. long, and 5·6 per cent faster than a line $4\frac{1}{2}$ in. long (Tinker and Paterson 1929). However, when a line of 9-point Scotch Roman 3 in. long with 2-point leading was presented first as the standard, a similar line of the same length without leading was read only 1·7 per cent faster than a line without leading of $2\frac{1}{4}$ in., and only 2·5 per cent faster than a line of 5 in. (Tinker and Paterson 1929). This discrepancy was clearly not due simply to the slightly smaller size of type in the second experiment. It could presumably have been prevented by an experimental design in which one condition was presented to each subject (Poulton 1959 a, p. 7).

3. A third difficulty is in the interpretation of experiments which do not yield statistically reliable differences. Negative results do not necessarily mean that differences do not exist. The design of the experiment may not

have been sufficiently sensitive to reveal real differences which were not very large. This criticism applies to Burt as well as to Tinker and Paterson. In Burt's experiments difficulty in reading was revealed both in the rate of reading and in comprehension as measured by questions answered directly afterwards (1959, p. 5). Neither measure alone gave reliable differences with adults, and Burt does not appear to have combined the two measures from each adult into a single measure of the *rate of comprehension* (score for comprehension divided by time for reading).

Tinker and Paterson simply demanded a certain limited level of comprehension from their readers by inserting in the second half of each paragraph one inappropriate word which had to be underlined. They measured rate of reading. This technique ensured that when reading became difficult, comprehension did not fall below the level required to pick out the incongruous word. Beyond this point difficulty was reflected entirely in slower reading. But there must have been a substantial range above this minimal level over which comprehension could vary without affecting Tinker and Paterson's measure of difficulty. Their experiments were thus not as sensitive as they could have been if comprehension had been held completely constant. By holding time for reading constant and measuring changes in comprehension, a recent experiment on scientists (Poulton 1959 a and b) has shown a statistically significant difference between a modern style and a revised old style of type. The fact that previous experimenters have found no significant differences in the rate of comprehension between the modern and old styles may thus have been due to the lower sensitivity of their experimental designs.

4. A fourth difficulty in interpreting experimental results is that the variables *interact* with each other : it is not possible to predict the effect of simultaneous changes in several different dimensions from the effects of each individual change carried out separately. For example, large type is read more easily when printed in long lines, but small type is read more easily printed in short lines (Tinker and Paterson 1931). Also an unfavourable change in a single variable may not yield an experimental difference which is detectable statistically, whereas the combination of two or more unfavourable changes may be detectable (Tinker 1948 b).

Until more experiments have been carried out it is clear that printers, especially in this country, cannot be sure that they are printing in the style which is the most helpful to their readers. What appears to be needed now is fairly specific research comparing styles of print which are practical alternatives for a particular purpose upon a sample of the population concerned (e.g. Poulton 1959 a). Additional basic work varying one parameter at a time does not appear to be required. Enough of this has already been carried out by Tinker and Paterson to give a fair idea of the likely effect of single isolated changes. What cannot be predicted for certain is the effects of a number of simultaneous changes.

When more research for special purposes has been carried out, it may be possible to decide whether additional basic work is required. If so it will probably be of a rather different kind from most previous research in this field, in that it will be concerned with the interactions of specific variables, including as one the type of reader.

Il existe une relation entre les caractères d'imprimerie et la vitesse de compréhension du lecteur ; cet article réunit et évalue des données expérimentales relatives à ce sujet. Des difficultés d'interprétation sont apparues et on suggère le genre de recherche dont à présent le besoin se fait le plus sentir.

Diese Arbeit sammelt und bewertet die vorliegenden experimentellen Befunde über die Beziehungen zwischen Druck-Stil und Auffassungsgeschwindigkeit beim Lesen. Es wird darauf hingewiesen, dass die Befunde schwer zu deuten sind und vorgeschlagen, welche Untersuchungsart am dringlichsten wäre.

References

Burt, C., 1959, *A Psychological Study of Typography* (Cambridge : University Press).

Burt, C., Cooper, W. F., and Martin, J. L., 1955, A Psychological study of typography. *Brit. J. Statist. Psychol.*, **8**, 29–58.

Luckiesh, M., and Moss, F. K., 1942, *Reading as a Visual Task* (New York : Van Nostrand).

Paterson, D. G., and Tinker, M. A., 1940, *How to Make Type Readable* (New York : Harper).

Poulton, E. C., 1959 a, Effects of printing types and formats on the comprehension of scientific journals. *Med. Res. Co. Appl. Psychol. Res. Unit Rep.* No. 346, Cambridge* ; 1959 b, Effects of printing types and formats on the comprehension of scientific journals. *Nature*, **184**, 1824–1825.

Tinker, M. A., 1948 a, Readability of book print and newsprint in terms of blink-rate. *J. Educ. Psychol.*, **39**, 35–39 ; 1948 b, Cumulative effect of marginal conditions upon rate of perception in reading. *J. Appl. Psychol.*, **32**, 537–540 ; 1960, *Legibility of Print* (Ames, Iowa : Iowa State College Press).

Tinker, M. A., and Paterson, D. G., 1929, Studies of typographical factors influencing speed of reading. III—Length of line. *J. Appl. Psychol.*, **13**, 205—219 ; 1931, V—Simultaneous variation of type size and line length. *J. Appl. Psychol.*, **15**, 72–78 ; 1949, Speed of reading nine-point type in relation to line width and leading. *J. Appl. Psychol.*, **33**, 81–82.

* Available from the author.

AN INVESTIGATION OF PILOT SKILL IN AN INSTRUMENT FLYING TASK

By D. C. V. Simmonds

Medical Research Council, Applied Psychology Research Unit, Cambridge

Seventeen pilots of varying degrees of experience each performed two complex manoeuvres on each of two separated occasions. The differences between their performances on the two trials were measured, and it appeared that the more experienced pilots showed smaller changes between the two trials than did the less experienced pilots. The effect of experience on *accuracy* in carrying out manoeuvres was less marked.

These results suggest that measurement of *consistency* may be useful as an objective index of pilot performance for comparing training methods and environmental conditions.

§ 1. Introduction

This experiment was intended to investigate the possibility of using *consistency* as a measure of skill (Gerhard 1954 ; Lewis 1956) in relation to an aircraft pilot's performance.

In the selection of an experimental task consideration was given to a study of pilot performance during approach and landing. However, early trials showed that accurate recording of position on the approach path was not possible with any method available at the airfield at which the work was to be carried out. Accordingly, an instrument flying task was chosen.

§ 2. Method

Flying tests were carried out in a Chipmunk aircraft, the subjects flying the machine from the front seat. The subject's helmet was fitted with a visor, permitting normal viewing of the cockpit interior, but preventing reference to external objects. The rear cockpit was occupied by the experimenter, who, while acting as safety pilot, operated the remote controls of a camera which photographed a panel of instruments situated in the fuselage immediately aft of the rear seat. The 'automatic observer' panel consisted of airspeed indicator, direction indicator, altimeter and engine r.p.m. indicator, as on the standard panel, and in addition a vertical accelerometer, elevator angle indicator, and a clock. The panel was photographed twice a second by a Bell and Howell A-4 camera, spring-driven, and operated by a solenoid. The instruments and camera were not accessible during flight, and thus a recording time of approximately 4 minutes was available on each flight.

The task was divided into two sections :

(*a*) Fly straight and level for one minute at cruising airspeed (85 knots), at the end of the minute carry out a climbing turn to the left at standard rate (3° per second) gaining 500 feet in height during the minute occupied by the turn. A reduction of airspeed to the normal climbing speed of 70 knots is to be established at the beginning of the turn.

(*b*) Fly straight and level for one minute at 85 knots, at the end of the minute carry out a standard rate descending turn to the right at 70 knots, losing 500 feet in height during the minute.

Before flight the subject was briefed with the aid of the diagram, shown in Fig. 1. When airborne, he was asked to fly on instruments to a height suggested by the experimenter—usually 2500 or 3000 feet—a starting height which would put the climb well within the range of the aircraft's performance. Variations were made on occasions to avoid excessively turbulent atmospheric conditions. Stop watches were synchronized in each cockpit, and, when the subject had settled down, he was told to carry out (*a*), followed by (*b*) after an interval of one minute, until the manoeuvres had been performed four times. Recordings were taken during the fourth set of manoeuvres of the last 30 seconds of the straight and level sections and the turns. The subject was unaware when recording was taking place. The entire procedure was repeated in a second flight on another day.

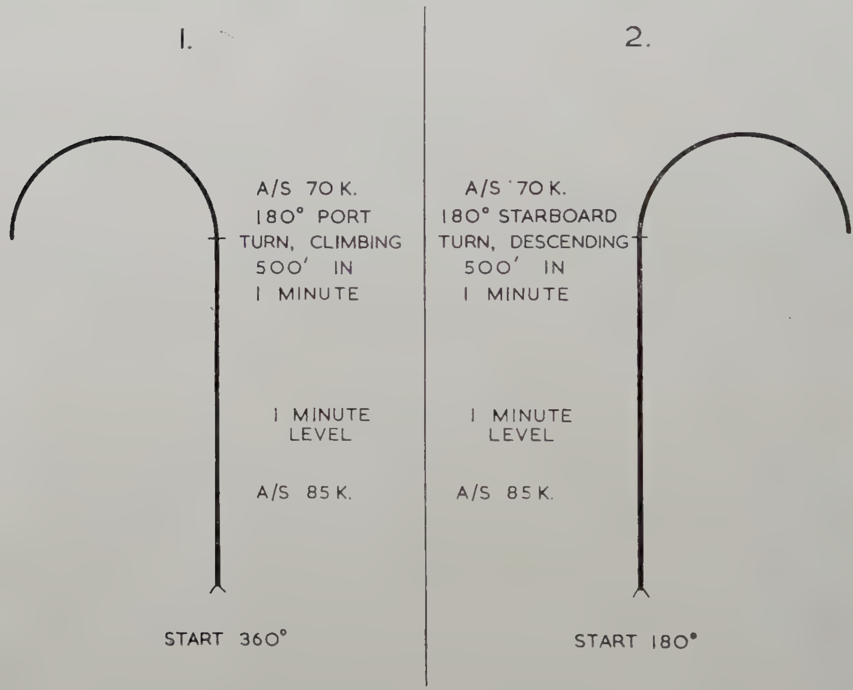

Figure 1. Diagram of the Test Manoeuvres.

The obtaining of subjects illustrated strikingly the practical difficulties that attach to a field study of this kind. It was found to be, for one reason or another extremely difficult to obtain the same subjects on two occasions separated by an appreciable interval of time.

The subjects on whom the final comparisons were based are shown in Table 1.

§ 3. Results

Each subject's performance was analysed in terms of deviation from the perfect pattern, readings from the airspeed indicator, direction indicator and altimeter being plotted at 10-second intervals. For each instrument the sum of these deviations is referred to below as the absolute error score.

Table 1. Particulars of Subjects

	No. in Group	Mean Age (years)	Mean Flying Hours	Instrument Flying Experience (hours)	Flying Time between first and second tests	Instrument Flying Time between tests
1. Cambridge University Air Squadron Students	3	20	97	11	$3\frac{1}{2}$ hrs	30 mins
2. Students from R.A.F. No. 3 Flying Training School	7	21	144*	11	44 hrs	$3\frac{1}{4}$ hrs
3. Qualified R.A.F. Pilots	4	27	854	97	$5\frac{1}{4}$ hrs	70 mins
4. Cambridge University Air Squadron Instructors	3	28	1604	213	3 hrs	40 mins

* Includes pre-service flying.

Engine r.p.m. settings were also plotted. The elevator angle indicator showed to some extent whether the pilot was making an abnormal number of adjustments to attain his object. Further information about an unexpected performance shown by the plots of the readings of the conventional instruments could be obtained by examining the indications of the vertical accelerometer and elevator angle indicator at closer intervals.

Consistency scores were obtained by considering separately for each instrument the two identical manoeuvres carried out by each subject on his two separate flights. The absolute difference between the instrument readings for each 10-second point was summed over the whole manoeuvre. Thus each subject had six consistency scores: one for the airspeed indicator, one for the direction indicator and one for the altimeter in the climb, and one for each of the three instruments in the descent.

3.1. *Consistency*

A multivariate analysis of the six consistency scores shows significant differences between the four groups ($P<\cdot 05$). In other words the groups did not give the same pattern of performance. If we rank the pilots for consistency on each instrument and then add the ranks for each pilot we find a significant rank correlation between this score and flying hours on this type of aircraft: $\tau=\cdot 51$, $P<\cdot 01$. In other words, more experienced pilots were more consistent.

The correlations for the different instrument were, however, far from equal, as shown in Table 2.

Table 2. Correlations (τ) between consistency scores and flying hours

	Climb	Descent
Airspeed	·29	·58
Altitude	·24	·02
Heading	·63	·08

3.2. *Accuracy*

The pilots were ranked again according to their Absolute Error Scores on each instrument and the ranks summed for each pilot. The pilots were then divided into those above and below the median on the consistency score and the absolute error score, to give Table 3. Absolute error was found to be correlated with consistency ($\tau=\cdot 40$, $P<\cdot 05$), but not, in these results, as closely to flying hours as was consistency (τ with flying hours$=\cdot 28$, $P>\cdot 05$).

Table 3. Relationship of consistency to accuracy

	Consistency	
Accuracy	High	Low
High	6	2
Low	2	7

3.3. *Experimenter's observations*

Observation during testing, suggested that Groups 3 and 4 were under considerably less strain in achieving their object than members of Groups 1 and 2. The less experienced students tended to concentrate on one instrument for too long a period, so that errors built up on other instruments. The more experienced pilots were, of course, better able to interpret the general situation from the instrument readings, but showed a wide diversity of ability to make any necessary corrections quickly. In most cases a considerable practice effect was observed, although it could not be recorded with the apparatus used.

§ 4. Discussion

From these results, consistency in flying performance shows some promise as an objective measurement of skill. Obvious applications of such a measure, if further developed, are as an index of the effectiveness of different training methods and as a sign of the effects of various stresses upon pilots. However, it would require considerable further study before being applied as a routine.

It would probably be economical to precede any further application of consistency measurement in the air by laboratory studies on the topic. Such experiments could suggest, for example, whether the reliability of a consistency measure can best be improved by taking a large number of simultaneous measures or by repeating a single measure on numerous occasions rather than only two.

The writer wishes to thank the Psychological Sub-Committee of the Flying Personnel Research Committee and its Chairman, Sir Frederic Bartlett, for providing facilities for this experiment. The Officers Commanding, staff and students of the various Flying Training Command units merit special mention for their cooperation.

Dr. Mervyn Stone gave most valuable statistical advice, and Mr. D. E. Broadbent's aid in the interpretation of the data was of inestimable value.

17 pilotes, inégalement expérimentés, ont accompli deux manoeuvres complexes dans deux occasions différentes. On a mesuré les différences entre leurs deux performances lors des deux essais, et il est apparu que les pilotes qui avaient plus d'expérience présentaient de moindres changements d'un essai à l'autre que ceux qui en avaient moins. L'effet de l'expérience sur la précision d'exécution des manoeuvres était moins prononcé.

Ces résultats suggèrent que la mesure de la stabilité de performance pourrait servir d'indice objectif de la performance de pilotage lorsqu'on compare des méthodes d'entraînement et l'influence des conditions ambiantes.

17 Piloten mit verschiedener Flugerfahrung führten bei zwei verschiedenen Gelegenheiten zwei komplette Flugmanöver im Instrumenten-Flug durch. Die Unterschiede der Leistungen in den beiden Versuchen wurden gemessen. Es schien, dass die erfahreneren Piloten geringere Unterschiede zwischen den beiden Versuchen zeigten als die weniger erfahrenen. Die Wirkung der Erfahrung auf die Genauigkeit bei der Durchführung eines Flugmanövers waren dagegen weniger deutlich.

Diese Ergebnisse machen es wahrscheinlich, dass die Messung der Gleichmässigkeit bei der Ausführung eines Flug-Manövers als objektiver Index für den Vergleich von Übung-Methoden und Umgebung-Bedingungen nützlich sein könnte.

References

Gerhard, D. J., An exploratory experiment in landing techniques. Unclassified, Flying Training Command Memo. No. 69. Flying Training Command Report No. 37 (F.P.R.C. 890), Aug. 1954.

Lewis, R. E. F., 1956, Consistency and car driving skill. *Brit. J. Industr. Med.*, **13,** 131–141.

THE DESIGN AND EVALUATION OF MAINTAINABLE PACKAGING METHODS FOR ELECTRONIC EQUIPMENT*

By J. M. McKendry, J. F. Corso[1] and G. Grant

HRB-Singer, Inc., State College, Pennsylvania

This study deals with the use of new packaging techniques to enhance the maintainability of electronic equipments. Three techniques were devised : (*a*) a ‘ component grouping ’ method, (*b*) a ‘ circuit grouping ’ method, and (*c*) a ‘ logical flow ’ method. Each of these methods was evaluated against the present ‘ standard ’ packaging method by constructing the same equipments by each of the four methods and determining the degree to which the performance of technicians was effected by troubleshooting each different type.

The experiment employed two skill levels of technicians : (*a*) those just completing training, and (*b*) those with more training and experience. In addition, two levels of equipment complexity were employed : (*a*) a simple radio receiver, and (*b*) a complex radar simulator. Finally, two main measures of performance were used : (*a*) the time taken to localize a fault to a component, and (*b*) the percentage of the total number of possible failures eliminated per three-minute time interval.

Consistent with the theoretical hypotheses advanced, one of the new methods, the logical flow technique, was found to be superior to the standard method in many cases. For example, a significant decrease in fault localization time (a 40 per cent saving) was achieved by the use of this new technique when experienced men worked on complex equipment (the radar simulator). In addition, significant increases in the percentage of failure possibilities eliminated were gained by the use of the new method in the following cases : (*a*) when inexperienced men serviced simple equipment, (*b*) when experienced men located hard troubles in simple equipment, and (*c*) when experienced men serviced complex equipment.

The other new methods, the components and circuit methods, were found to be superior to the standard method, but their superiority was not as marked as that demonstrated by the logical flow technique. The advantages to be gained by use of the new methods easily seem to outweigh problems met in implementing a new construction policy based upon them.

§ 1. Introduction

It is now generally known that the costs of equipment maintenance greatly exceed the initial purchase price for any given electronic device (Runyon, 1957). Several attempts have been made to devise a set of rules by which more ‘ maintainable ’ equipment might be constructed. The first rules devised were concerned with the exterior or more mechanical features of electronic equipment (Folley and Altman 1958). However, these rules, while appropriate, were too restricted to alter substantially the problems confronted in the maintenance of complex operational systems.

In the course of developing a handbook aimed specifically at increasing the ‘ maintainability ’ of electronic equipment (McKendry, Grant, Corso and Brubaker 1959), the authors reached the conclusion that it should be possible

* This research was supported by the U.S. Navy under contract N 61–339–330, monitored by the Flight Trainers' Division, Human Engineering Department of the U.S. Naval Training Device Center, Port Washington, N.Y. The opinions and conclusions in this report are those of the authors and are not to be construed as reflecting the views or endorsement of the Department of the Navy. The authors would like to acknowledge the help of George Herlt who acted as a special engineering design consultant for equipment fabrication. Both he and Fred Scheihing were concerned with the difficult task of translating the packaging philosophies into actual hardware. Additional assistance was rendered by Mr. Karl Goosman and Mrs. Margaret McKendry.

[1] J. F. Corso is also a Professor of Psychology at Pennsylvania State University.

to attack the problem of maintenance at a more basic level by designing and evaluating new packaging methods, which would allow equipment to be assembled in such a way that unique physical appearances would provide the repairman with enough information to decrease the difficulty of trouble-shooting.

This article summarizes the results of an experimental study performed to develop and evaluate several new methods of packaging electronic equipment. The evaluation of the new methods was based upon comparisons with standard methods in terms of (*a*) the reduction in time of fault localization, and (*b*) the amount of additional information gained as a function of specific troubleshooting tests.

§ 2. Experimental Design

The experimental plan involved three main variables : the methods of packaging ; the complexity of equipment ; and the skill level of the repairman.

Two pieces of equipment were constructed according to each of four packaging techniques : the first was a radio receiver, and the second a radar simulator. The receiver was selected to represent a ' simple ' device while the simulator was considered as a piece of ' complex ' equipment. A detailed description of these experimental equipments may be found elsewhere (McKendry, Corso and Grant 1959).

Subjects of two levels of technical skill were used : (*a*) 27 Navy repairmen who had completed only the basic U.S. Navy electronic course and who had had no previous practical experience ; and (*b*) 22 Navy repairmen with more advanced schooling and at least one year's work experience.

The inexperienced technicians were tested only on the simple equipment since spot checks showed that there was little likelihood of their solving the problems built into the complex equipment. The experienced repairmen were tested on both pieces of equipment.

2.1. *Explanation of Packaging Technique*

In the ' *component packaging* ' method all similar components are grouped in one place on the equipment. For example, all tubes are together, all transformers, etc. Inexpensive components, such as resistors and condensers, are placed on separate plug-in type boards mounted underneath the chassis. Resistors and condensers are subdivided on different boards. For resistors, there is a ground board, a feedback board, a B+ board, etc. This permits the resistors to be checked by a low-cost mass replacement technique. In addition, the technique of component grouping enhances the possibility for the development of simplified checkout procedures. A photograph of the ' component packaging' technique as applied to the radio receiver is given in Figs. 1*a* and 1*b*.

In the ' *circuit packaging* ' method, all similar or identical circuits are physically grouped together. Each circuit is placed in a separate module, while the tube associated with the circuit is placed on top of the module. Since most equipment can be considered to be a series of functioning circuits, it seems reasonable to suppose that troubleshooting would be aided by using a technique of circuit grouping. A photograph of the receiver built by the ' circuit packaging ' technique is given in Figs. 2*a* and 2.*b*

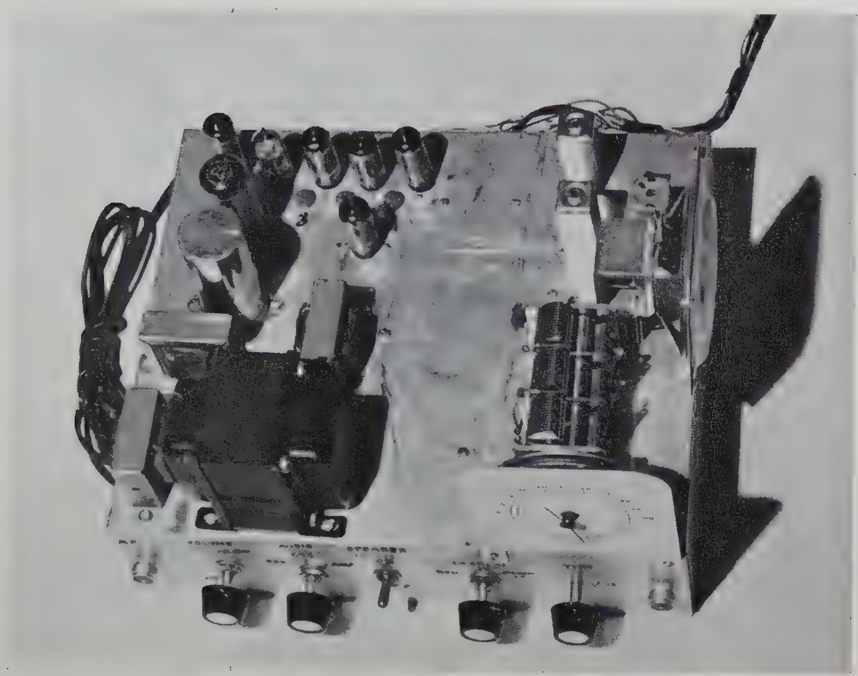

Figure 1 (*a*). Top view of receiver packaged using the component grouping method.

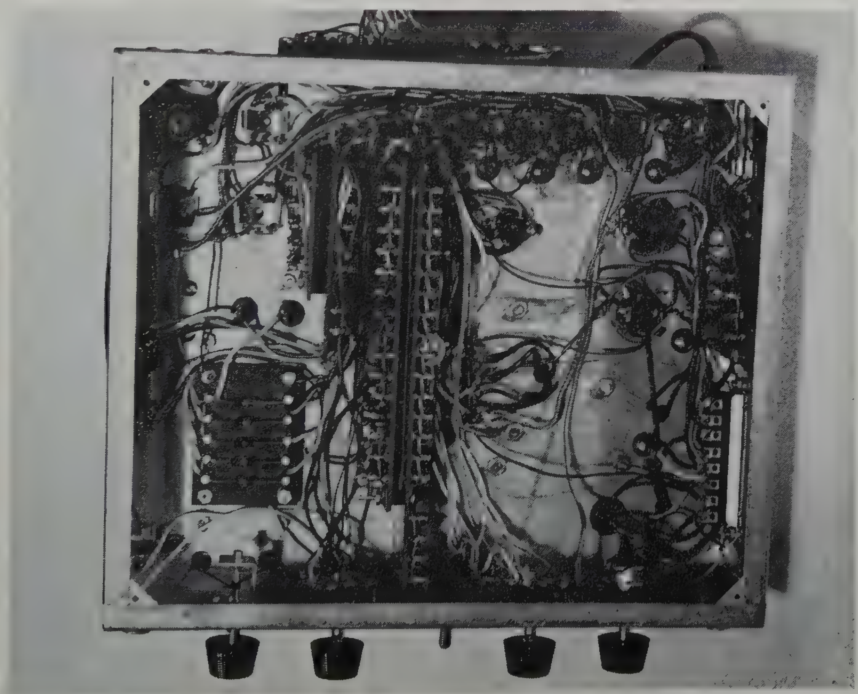

Figure 1 (*b*). Bottom view of receiver packaged using the component grouping method.

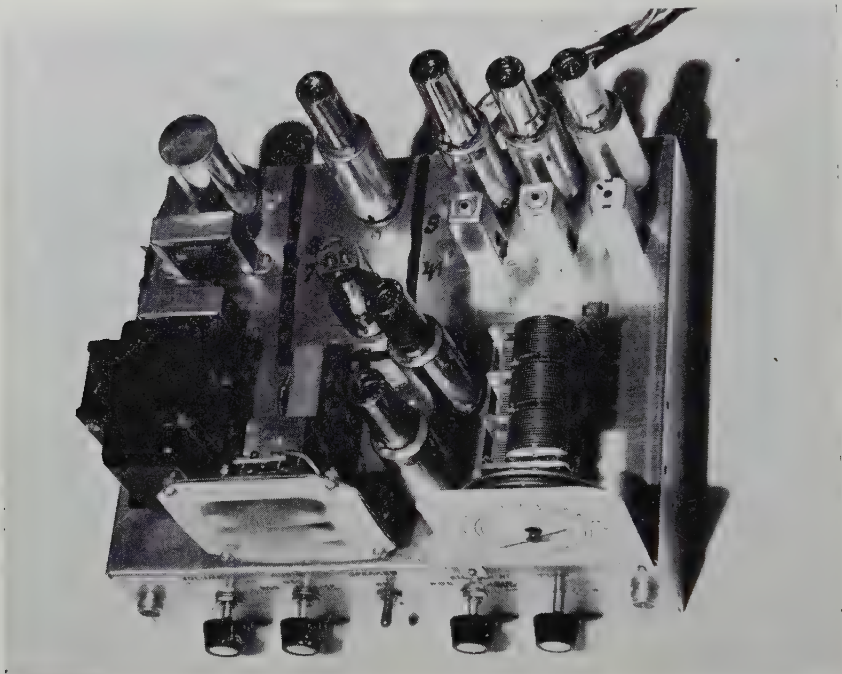

Figure 2 (*a*). Top view of receiver packaged using the circuit grouping method.

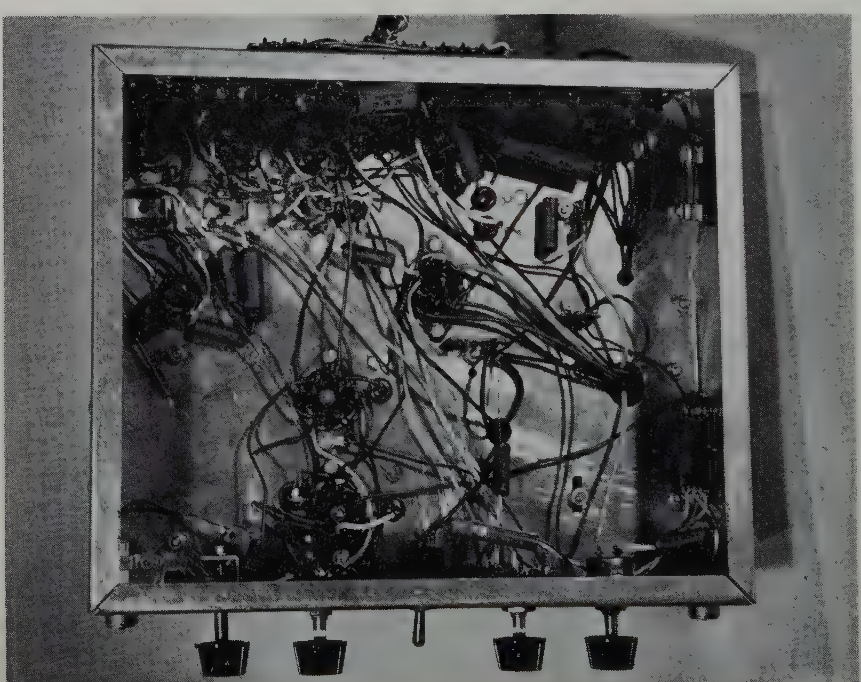

Figure 2 (*b*). Bottom view of receiver packaged using the circuit grouping method.

The last new method, '*logical flow packaging*', represents a combination of two approaches. The first of these is the suggestion that the use of modules and subchassis will enhance maintenance (Folley and Altman 1958). The second approach utilizes the concept of 'data flow' or 'signal flow' which has been shown to be critical to the maintenance problem (Rulon, Schweiker and Gilbert 1958, and Miller and Stebodnick 1958). The 'logical flow packaging', therefore ; incorporates two main features : (*a*) the use of modules and subassemblies so selected that only a single, simple input and output check will be necessary to isolate a trouble to that unit, and (*b*) a clear indication of uni-directional signal flow within the given piece of equipment. An example of the 'logical flow' technique is shown for the superheterodyne receiver in Figs. 3*a* and 3*b*.

To evaluate each of the three experimental packaging techniques, comparison measures were obtained on the radio receiver built according to the '*standard packaging*' method (see Fig. 4). It does not employ modules, nor does it attempt to keep a straight-line representation of signal flow. The parts on top of the chassis are appropriately spaced to distribute weight and heat. No consistent grouping-technique is used in the receiver.

For the purposes of this study, the test jacks on all experimental equipments were mounted on the outside of the chassis and were numbered for cross-referencing with schematic diagrams. The same test points were used in all equipments.

2.2. *The Task*

The subject's task for all experimental conditions was to isolate a fault to a given component. Each subject acted as his own control, i.e. he was tested on the 'standard' equipment as well as on each of the three experimental models. Each subject received two troubles under each 'condition' or method of packaging. This required a total of eight troubles for both the simple and complex equipments since the subjects were not permitted to be tested twice on the same problem. A trouble panel was devised for each equipment so that any one of the eight troubles could be inserted by the experimenter simply by throwing a switch.

In general the problems involved faulty tubes, resistors, transformers, and chokes. An exact listing of the specific troubles introduced in the experimental sessions has been provided in more detail elsewhere (McKendry, Corso and Grant 1959).

When working on the receiver, the subject was allowed 30 minutes per trouble. When working on the simulator, he was given 40 minutes per trouble. The time was recorded by an observer with a stop watch.

For experimental purposes, it was arbitrarily decided to start each technician with a properly functioning piece of equipment. A fault was then placed into the equipment by throwing the switch. The subject was informed that a trouble now existed and that he had so much time to find it.

If the subject decided to make a replacement, he informed the experimenter that he wished to replace a certain module or component. The experimenter noted the simulated change and informed the subject whether or not the trouble had been corrected. For module replacements, the subject was also required to locate the faulty component. By this technique only the time of

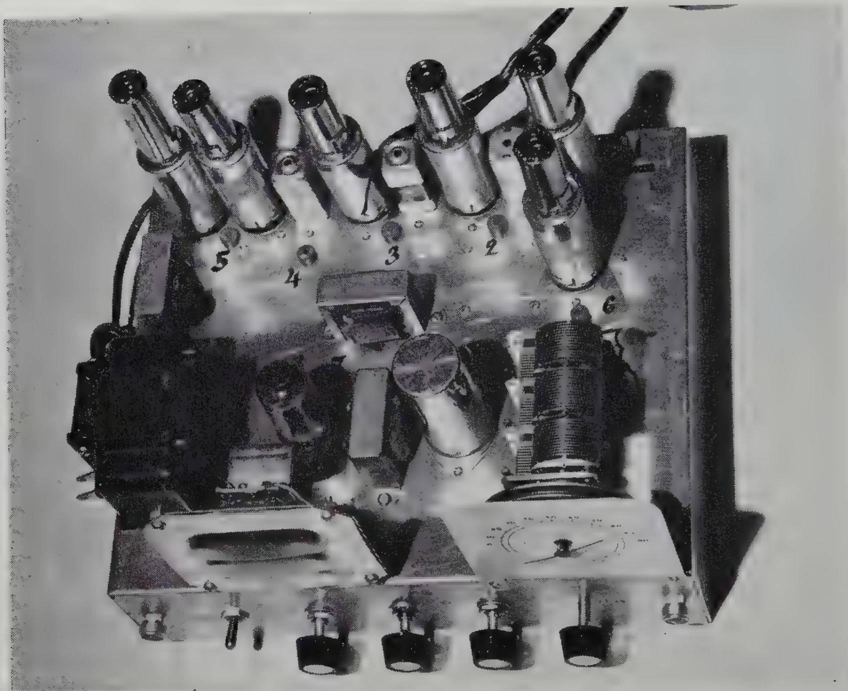

Figure 3 (*a*). Top view of receiver packaged using the logical flow method.

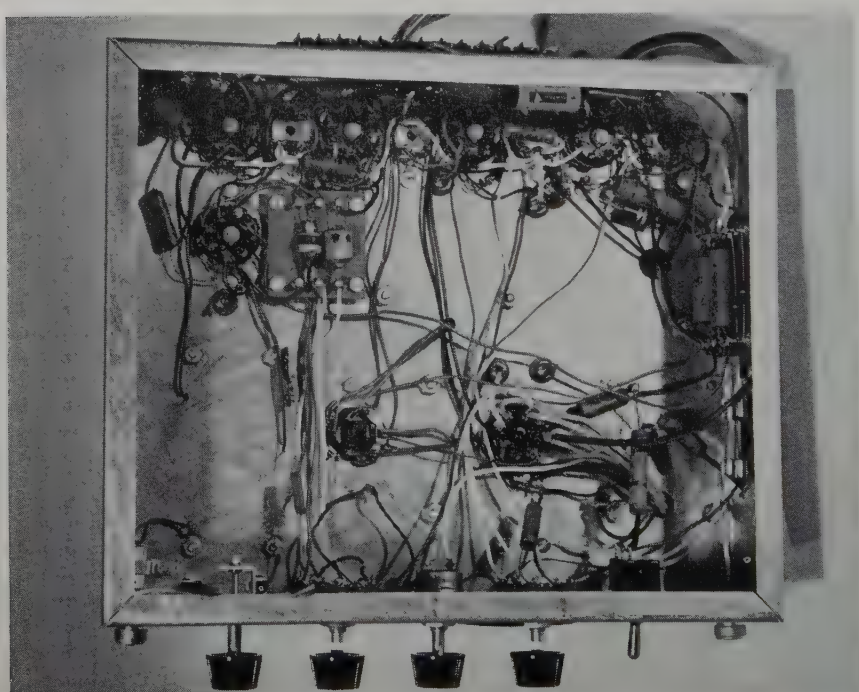

Figure 3 (*b*). Bottom view of receiver packaged using the logical flow method.

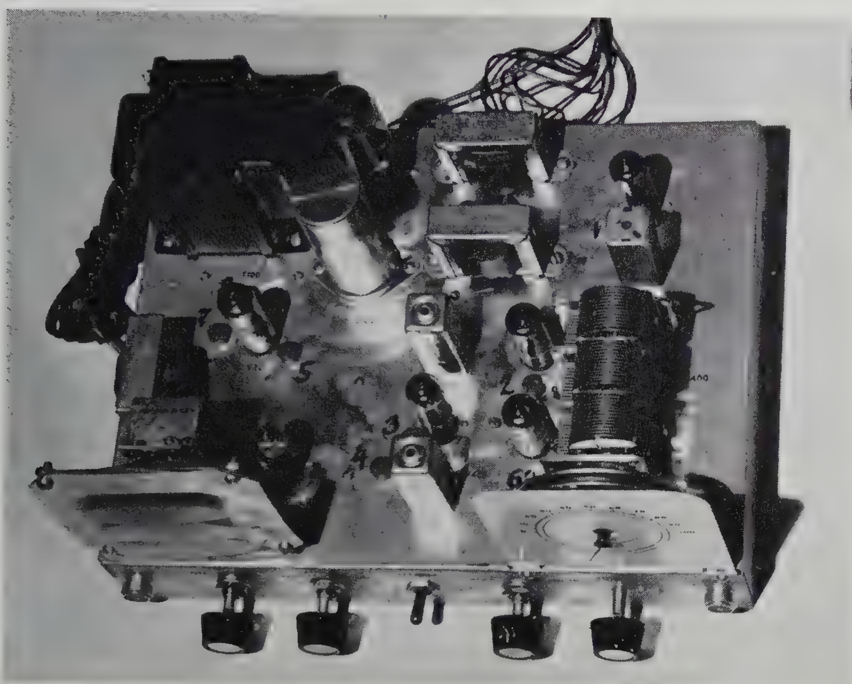

Figure 4 (*a*). Top view of receiver packaged using the standard method.

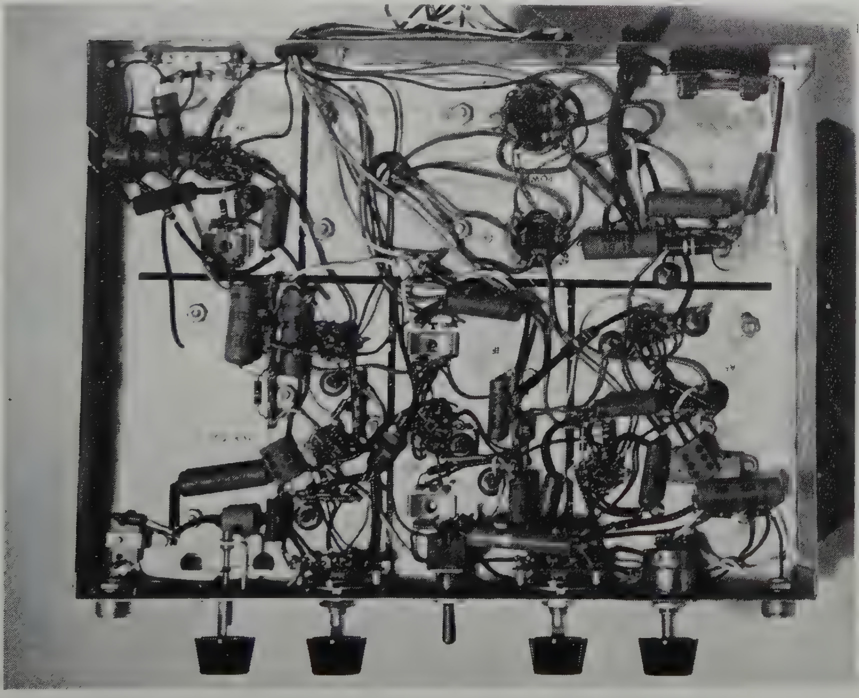

Figure 4 (*b*). Bottom view of receiver packaged using the standard method.

fault localization was recorded. The subjects were not allowed to stumble upon the correct replacement by making a series of random guesses.

2.3. *Performance Measures*

Although the maintenance process can be divided into the stages of fault detection, fault localization, and parts replacement and/or repair, the criteria of the present study were restricted to the stage of fault localization. The first performance measure, fault localization time, was operationally defined as the time taken to isolate successfully a trouble to a single faulty component once its presence had been detected. The second measure was an estimate of the amount of troubleshooting information that the technician obtained per unit of time. This information was equated to the percentage of components or segments that he eliminated as the possible source of the malfunction during successive time periods. This measure was included for two reasons : (*a*) to give a more complete indication of the relative efficiency of the packaging techniques in terms of their effects on troubleshooting strategies, and (*b*) to compensate for the arbitrary maximal time limit assigned to the problems on each of the experimental equipments.

The second measure was obtained in the testing sessions by asking each subject at three-minute intervals whether he knew where the trouble was or was not. The answers to the question revealed whether the subject was getting any information from the checks he was performing. For example, when the subject replied that ‘ the power supply is all right ’, he was given credit for eliminating 14 potential component troubles. He was given no credit when his elimination was incorrect and he was never told whether or not his guess was correct.

In addition to the objective measures of performance, the subjects were asked which method was most preferred.

2.4. *Instructions*

An introductory one-hour class was held for each experimental group. All lectures were recorded on tapes so that each group would be given a standard set of instructions. In addition, each class was given a a copy of the lesson plan which presented a verbatim account of the material recorded on the tapes. The instructor referred the class to specific diagrams provided in the lesson plans, while the taped lecture explained them.

Before each lecture, a short introduction was given to each group. The purpose of this was to tell the subjects what was expected of them, why the class was taped, and why it was important to follow closely the instructions given. The general nature of the research was also mentioned. An announcement was made that a paper-and-pencil test would be given at the end of the lecture. The scores obtained on this test were later correlated with the time taken to locate a fault in a real troubleshooting problem situation.

The content of the course of instruction was carefully chosen so that no experimental packaging technique was specifically discussed. No block diagrams were provided since this would tend to structure the technicians’ thinking toward certain methods. Copies of the lesson plans and tests are reproduced elsewhere (McKendry, Corso and Grant 1959).

Each subject was tested within one to three days after the class instruction had been received. At the beginning of each testing period, each subject was given a set of general instructions which told him what to expect at each work station.

2.5. *Workspace Arrangement*

The subjects were tested in a large room at the Naval Air Station. Two standard work benches were placed together to provide adequate work space. Each work area was provided with the equipment to be serviced, an oscilloscope, a multimeter, a schematic diagram with test points indicated, a two-paragraph statement explaining the method of packaging being used, a copy of the lesson plan which explained the device being serviced, a wiring diagram of the modules (where necessary), a location chart, and a list of test points which included the proper resistance and voltage readings together with the waveforms which should be available at that point.

§ 3. Results

3.1. *Differences in Fault Localization Time*

An analysis of variance was used to evaluate the differences in fault localization time. Differences between any specific methods were evaluated by use of Tukey's tests for significant gaps and straggling means (Edwards 1954). Whenever significant results were noted using the F test, two non-parametric tests were also used for additional confirmation : the Friedman two-way analysis of variance (Siegel 1956), and the Jonckheere test (Jonckheere 1954).

Four parametric tests of significance were computed to evaluate the differences in fault localization time for the four different packaging methods. The first three tests involved the differences between experienced subjects : (*a*) on the easy receiver troubles, (*b*) on the difficult receiver troubles, and (*c*) on the radar simulator troubles. The fourth test evaluated the performance of inexperienced subjects for all troubles on the receiver. Of the four analyses, only one showed significant differences between packing methods. This was for the experienced personnel working on the complex (radar simulator) equipment. The results of this analysis are summarized in Table 1. In this case, the use of Tukey's tests showed that all three of the new methods were significantly better than the standard method. Although the logical flow method was best, it was not significantly better than the other two new methods.

Source	SS	df	ms	F	P
Packaging Methods	2,956·33	3	985·44	8·246	$<$·01
Subjects	6,037·88	21	287·52	2·406	$<$·01
Residual	7,528·34	63	119·50		
Total	16,522·55	87			

Table 1. Results of analysis of packaging techniques using rated (experienced) subjects on the radar simulator (complex equipment). Because of the demonstrated robustness of the F test (Lindquist 1954; Ray 1960) this analysis appears appropriate despite the fact that the distribution of fault localization times was positively skewed. Use of a non-parametric test, the Frieman two-way analysis of variance, confirmed these results by also showing the differences in packaging methods to be significant (P $<$.02).

The use of non-parametric tests on the same data led to the same conclusions. The Friedman test clearly showed that the mean performance ranks differed significantly ($P < \cdot 02$). This difference was described by using the Jonckheere test to test the hypothesis that localization time would be least for the ‘ logical flow ’ equipment ,greatest for the ‘ standard equipment ’ and intermediate for the ‘ circuit ’ and ‘ component ’ equipments. It was predicted that the time for the ‘ circuit ’ equipment would be less than the time taken to locate troubles on the ‘ component ’ equipment. This hypothesis was clearly confirmed ($P < \cdot 001$).

The implications of the significant differences noted for experienced men working on complex equipment for the fault localization criterion can be understood by noting how much localization time is saved by using these methods instead of the standard packaging techniques. There was a 40 per cent saving for the logical flow method, a 29 per cent saving for the circuit method, and a 26 per cent saving for the component method. The three other analyses which were conducted showed the same type of time savings, but these were not great enough to achieve statistical significance (actual values falling between 0–19 per cent.)

3.2. *The Percentage of Components successfully eliminated per given Time Interval*

The second criterion against which the various packaging techniques were evaluated was the percentage of components successfully eliminated by the subjects during each three-minute time interval for each problem. The scores were obtained by first counting the total number of components which were present in the entire receiver or simulator. By a similar procedure, the number of components in each section, stage, module, and board was determined. The subject's score at the end of each three-minute time interval was taken as the cumulative number of components eliminated up to and including that point in time. This score was expressed as a percentage by dividing the cumulative total by the total number of components in the equipment and multiplying by 100.

The results of the analysis of these data are given in Figs. 5–8. These figures provide a graphic representation of the cumulative percentage of eliminations obtained under each condition as a function of troubleshooting time. In each figure, the coefficient of concordance, W, is given to show the consistency of the ranking of packaging methods for all ten-time intervals. Notice that Figs. 5, 7, and 8 show a fairly high degree of consistency, but Fig. 6 indicates that some marked changes occur in the selective ordering of the four methods during the ten-time intervals. Nevertheless, at the end of the 30 minutes' testing period, the logical flow method shows the greatest percentage of eliminations in each case.

3.3. *Subjects' Impressions of the Best Method*

The data for this portion of the study were obtained from 12 inexperienced men and from 15 experienced men. These subjects were required to indicate which method of packaging they liked best. The percentage of men choosing each type differed greatly, depending upon the skill level. The following preferences and percentages were obtained from the inexperienced men :

(*a*) Standard Method—3 men or 25 per cent ; (*b*) Components Method—3 men or 25 per cent ; (*c*) Circuits Method—4 men or 33 per cent ; and (*d*) Logical Flow Method—2 men or 17 per cent.

The results of preferences stated by the experienced men were vastly different. All the 15 men questioned, preferred the logical flow method, with the circuits and components methods being mentioned once each for being as good as the logical flow method.

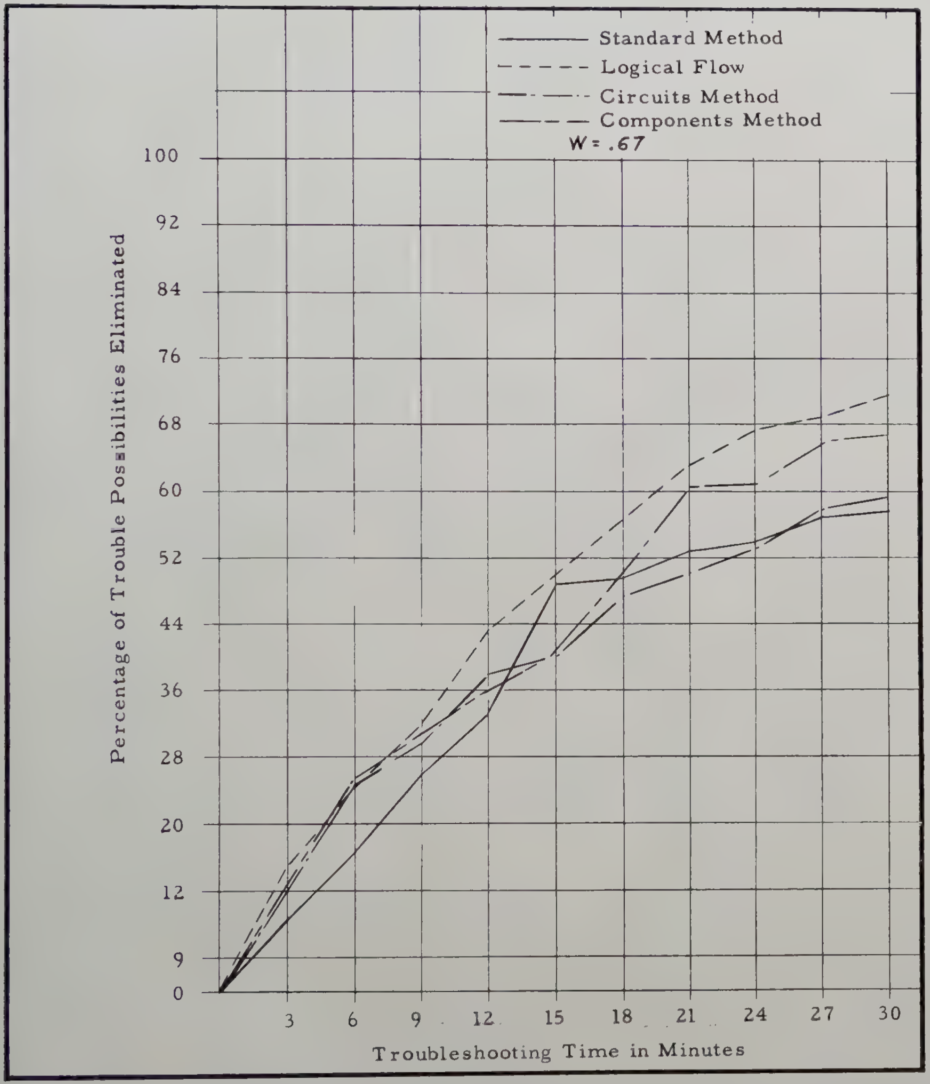

Figure 5. Cumulative performance curves of inexperienced men on receivers. The percentages were computed by dividing the number of components eliminated by the total number available.

3.4. *Relation between Troubleshooting and Paper-and-Pencil Test Performance*

The final analysis concerns the correlation of the paper-and-pencil test scores obtained at the end of the lecture session with the subject's average time taken in locating a fault in the troubleshooting session.

Three Pearson product moment correlation coefficients were computed : (*a*) for inexperienced men on the receiver, r = 0·15 ; (*b*) for experienced men on the receiver, r = ·06 ; and (*c*) for experienced men on the simulator, r = ·12. Since these coefficients indicate a low, positive correlation and confirm previous work (Stanler, Popham and Fattu 1956) they tend to raise doubts about the significance of evaluation procedures currently employed in numerous technical training schools.

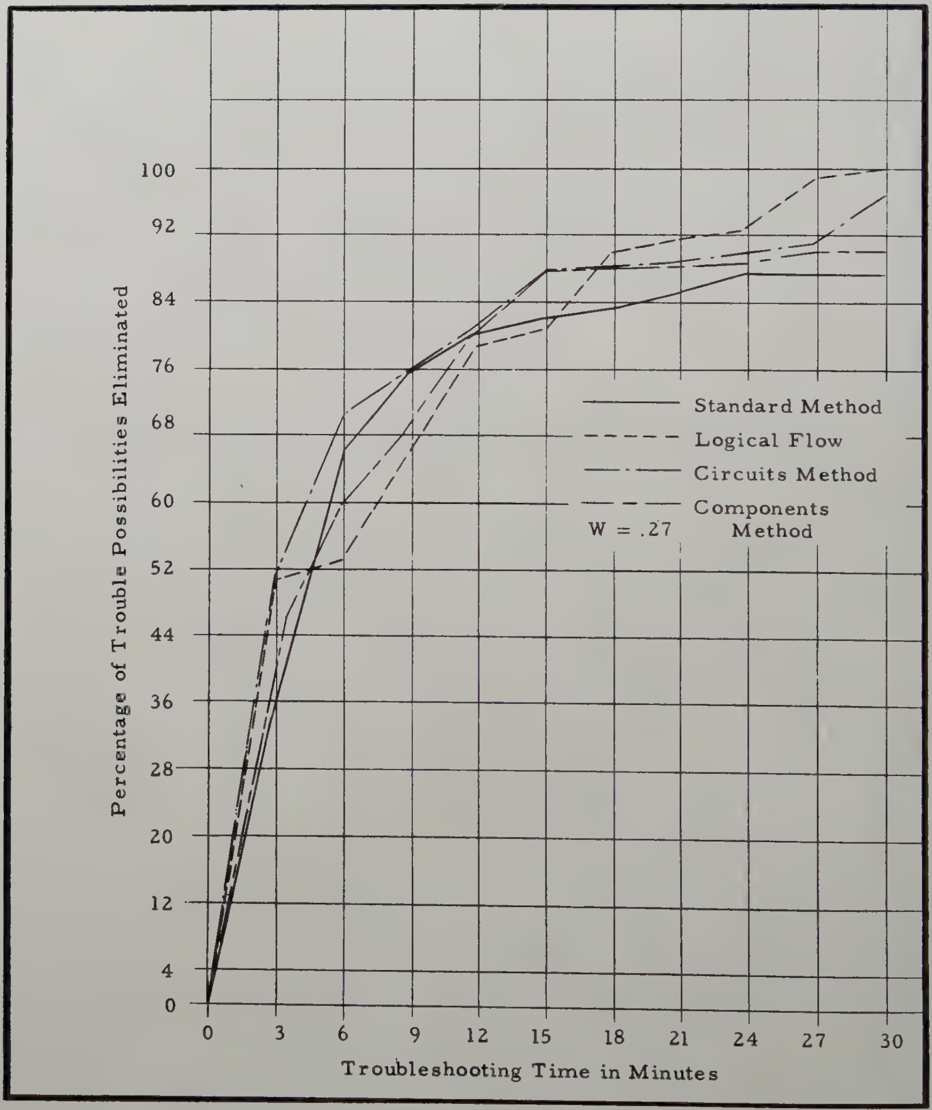

Figure 6. Cumulative performance curves of experienced men for easy receiver troubles.

§ 4. Discussion

The only significant difference was obtained for the experienced subjects working on the simulator equipment. The failure to obtain significant differences between packaging methods for the other three analyses could have

occurred as a result of any of the following : (*a*) criterion contamination provided by prior experience with the receiver, (*b*) inadequacies in the choice of the arbitrary time limit, and/or (*c*) the simplicity of the radio receiver which prohibited an adequate demonstration of the new packaging techniques.

While there is some evidence for each of these factors, it appears that the last point is most important. Notice that the advantages gained by the new methods in fault localization time seemed to be more marked with increasing task difficulty, i.e. were significant when working on difficult equipment. The same trend was illustrated even more strikingly when the second criterion, the information gained per given time interval, was used.

In Fig. 5 it can be seen that there are no great differences between packaging methods at any particular point in time. However, it is interesting to note that the ranking of conditions (*a*) logical flow, (*b*) components, (*c*) circuits, and (*d*) standard, was fairly consistent, $W = 0{\cdot}67$.

It should also be noted that the logical flow method was better than the standard method at all intervals and that the component method was better than the standard method at all times, except one. Only one experimental technique, the circuit method, did not clearly excel the standard method.

In interpreting the curves of Fig. 6, the frequent reversals indicated by the low coefficient of concordance, $W = 0{\cdot}27$, show poor consistency of the packaging effects over time. It may be reasonably concluded therefore that no differential packaging effects were demonstrated in this case.

Some consistent packaging effects seem to be noticeable in Fig. 7. It appears that the standard method was not as effective as the others in helping to eliminate receiver troubles. The coefficient of concordance is moderate, $W = 0{\cdot}63$, and shows a reasonable consistency among methods as a function of time. However, it should be noted that most of the inconsistency is due to the logical flow method which started slower than the component method in terms of eliminations, but improved at a steady rate until it became best. The logical flow technique showed similar trends in Figs. 5 and 6. The effects of the circuit and component methods do not seem to be as clear, except that they seem to fall somewhere between the standard and logical flow methods in terms of effectiveness. In general, they are much closer to the logical flow than to the the standard method. The standard method is consistently inferior to the other methods.

The performance of the experienced men on the simulators (see Fig. 8), quickly approaches a limit in a much shorter time in the logical flow method than the other methods. The curve for the standard technique rises slowly and asymptotes earlier. The curve for the circuit method seems to parallel closely the curve for the logical flow technique, with little difference appearing between them. Both curves seem to have some advantage over that for the component method, which in turn is better than that for the standard method. This relationship is highly consistent, $W = 0{\cdot}91$, with the only reversals occurring at the upper limit where scores were tightly grouped.

The acquisition curves of Fig. 8 can be smoothed to straight lines if the percentage of components eliminated is changed to a logarithmic function. When this is done, it becomes clear that the curves differ only in slope and not in shape. The significance of this finding is that use of the new methods does not change the type of learning even though it markedly changes the speed.

Examination of the third criterion, the analysis of the technicians' preferences, showed that no single method was especially preferred by a majority of inexperienced technicians. The results of the preferences stated by experienced men were vastly different. From the results, it appears that a significant advantage must accrue to a particular method before it becomes preferred. In addition, the finding suggests that technicians will probably offer little resistance to a change in packaging techniques, provided some benefits are gained in terms of maintenance requirements.

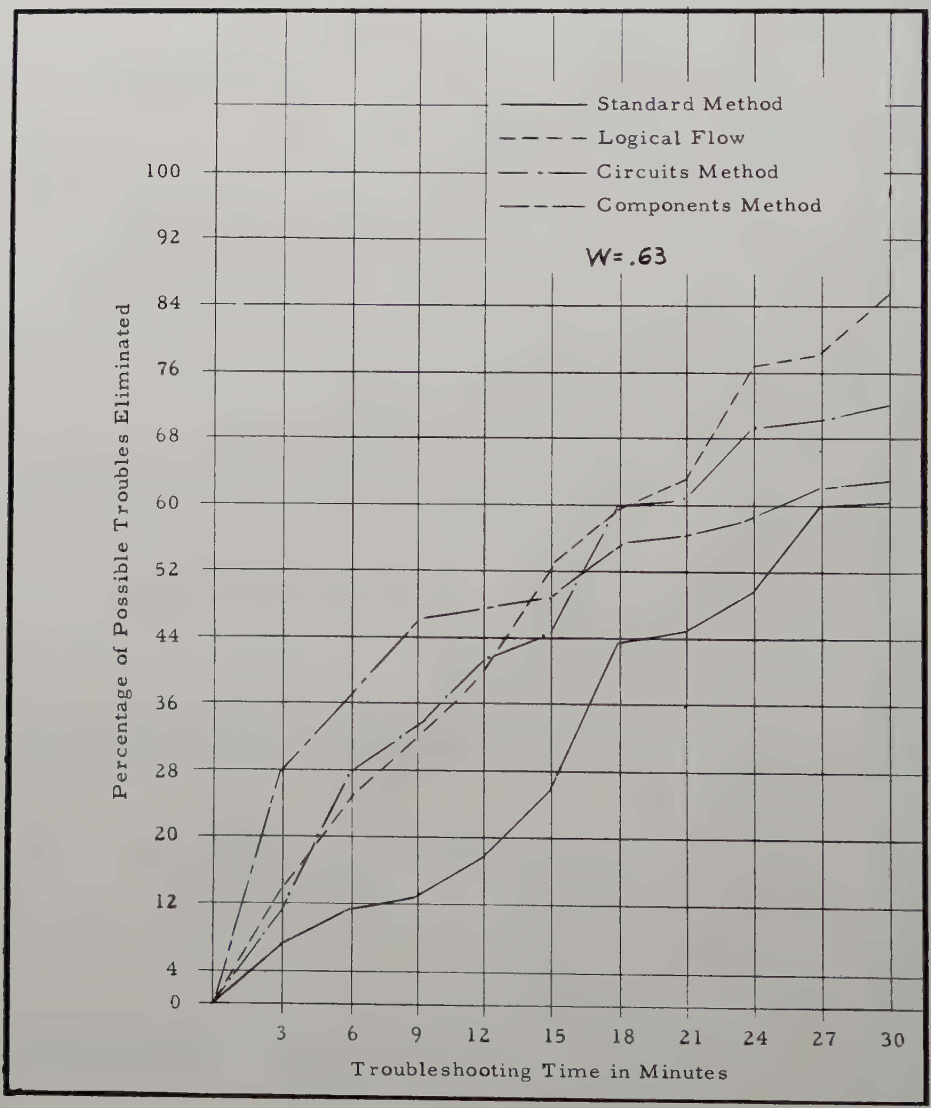

Figure 7. Cumulative performance curves of experienced men for hard receiver troubles.

The last data which described the relation of theoretically orientated paper-and-pencil tests of equipment understanding and measures of trouble-shooting proficiency would indicate that the two factors are almost independent.

It is therefore apparent that the use of the former measure in judging the latter is a practice having questionable value.

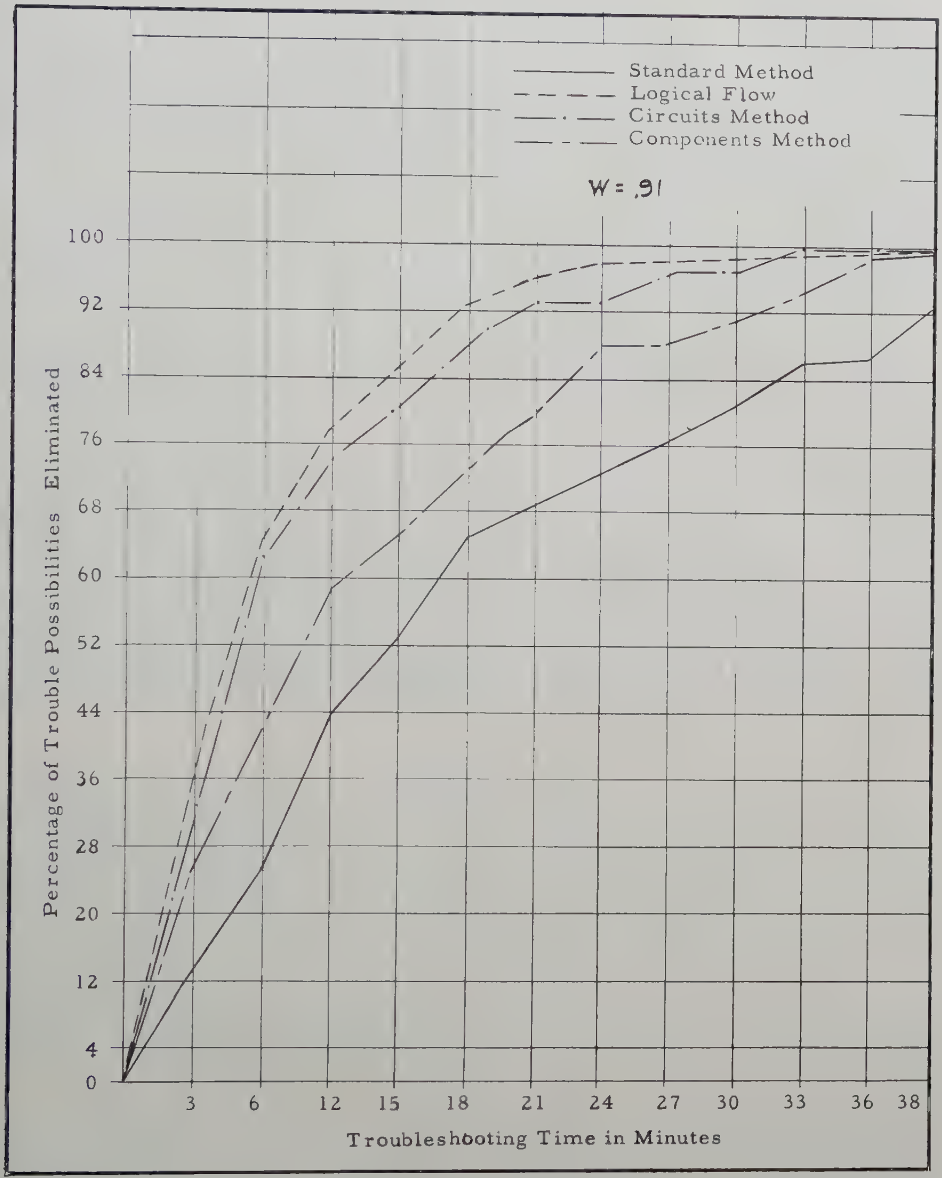

Figure 8. Cumulative performance curves of experienced men on the simulator.

Although prior experience in technical schools should have considerably favoured the standard method, the logical flow method seems to have much to recommend it. It was efficient enough to counteract the experience factor without the benefit of special training. If a small amount of training were provided in this method, the results would probably have been even more marked.

The logical flow method, if properly taught, should facilitate a more systematic method of troubleshooting than that currently available. In fact, in this study, the packages were constructed so that it would be easy for the technicians to use the half-split troubleshooting method (Miller, Folley and Smith 1953), which, according to a mathematical analysis, is the most efficient. Furthermore, the method could simplify the form of the supporting diagrams required in the troubleshooting of equipment. It could almost eliminate the need for the detailed schematics which have been demonstrated to cause considerable anxiety among inexperienced men working on complex equipment (Stanler, Popham and Fattu 1956).

Although all of these later claimed advantages are stated as hypotheses, it is felt that such performance advantages can be readily shown in subsequent research. The ability of the logical flow method to achieve a simplification of diagrams should be apparent, since the replaceable units on the chassis would correspond to the blocks on the flow diagram. With this technique, the troubleshooter should be able to isolate rapidly a trouble when he knows only in a rough fashion what is happening to the signal as it flows through the system.

Members of the engineering staff of HRB-Singer, Inc. have evaluated the new methods against the standard method in terms of five criteria : (*a*) electrical considerations, (*b*) mechanical problems, (*c*) reliability, (*d*) building time, and (*e*) ease of adaptibility for use with automatic checkout equipment. In each instance, the logical flow method was judged to be at least as efficient as the standard method. In (*e*) above, it clearly excelled the standard. The other methods, especially component packaging, have certain difficulties which can be overcome by following the suggestions presented in another report McKendry, Corso and Grant 1959).

Cette étude s'est occupée avec l'utilité de nouvelles techniques d'encaissement pour augmenter la facilité du maintien d'équipment électronique. On a inventé trois techniques : (*a*) une méthode de grouper les pièces (component grouping method), (*b*) une méthode de grouper les circuits (circuit grouping method), et (*c*) une méthode d'écoulement logique (logical flow method). On a comparé ces trois méthodes d'encaissement avec la méthode actuelle au moyen de construir la même espèce d'appareillage par chacune des quatre méthodes et ensuite d'établir le temps pris par les techniciens pour localiser les pannes dans les diverses espèces d'appareillage.

Les expérimenteurs ont employé des technicienc de deux niveaux de dextérité : (1) des personnes qui venaient compléter leur instruction, et (2) des personnes qui avaient plus d'instruction et plus de pratique. En outre, on a employé les appareils de deux niveaux de complexité : (1) un récepteur simple er (2) un 'simulateur' complex de radar. Enfin, on a employé deux mesures de performance : (1) combien de temps on a employé pour trouver une défaillance dans une partie, et (2) le pourcentage du numbre total des défaillances possibles qui on a pu éliminé par intervalles de trois minutes.

D'accord avec les hypothèses énoncés, on a découvert qu'une des nouvelles méthodes, celle d'écoulement logique, valait plus que la méthode actuelle. Par example, quand les techniciens expérimentés travaillaient avec les équipements complexes il y avait une diminution significative dans le temps employé pour localiser une panne, une diminution de 40%. En outre, l'emploi de al nouvelle méthode a apporté un augmentation significative en pourcentage de pièces bien functionnantes qu'on pouvait éliminer comme des sources possibles de défaillance dans les cases suivants : (1) quand les techniciens inexpérimentés cherchaient à localiser les défaillances dans les appareils simples, (2) quand les techniciens expérimentés cherchaient à localiser les défaillances plus difficiles dans les appareils simples, et (3) quand les techniciens expérimentés cherchaient à localiser les pannes dans les appareils compliqués.

Les autres nouvelles méthodes, la méthode de grouper les pièces et la méthode de grouper les circuits, valaient mieux que la méthode actuelle, mais leur superiorite n'était pas aussi évidente que celle de la méthode d'écoulement logique. Il semble que les avantages de faire usage des nouvelles méthodes supasseraient les problèmes qu'on rencontrefait en les employant.

Diese Abhandlung befasst sich mit dem Gebrauch neuer Zusammensetzungsmethoden, die die Erhaltung elektronischer Apparate vereinfachen. Drei Methode wurden erdacht :

(*a*) die 'Bestandteilgruppe-Methode'
(*b*) die 'Leitungengruppe-Methode'
(*c*) die 'rationale Impulsstrom-Methode'

Jede dieser Methoden wurde im Vergleich mit der bisherig normalen Zusammensetzungsmethode ausgewertet, indem man denselben Apparat auf jede der verschiedenen Methoden viermal zusammenstellte. Für jede Art Zusammensetzung wurde dann den Grad festgestellt, auf den das Verfahren der Techniker beeinflusst war, die Fehlereingrenzung ausüben sollten.

Die Techniker, die am Versuch teilnahmen, waren zweier Ausbildungsstufen : (*a*) einige, die eben aus dem ersten Ausbildungsprogramm kamen, und (*b*) andere, die weitere Ausbildung und Erfahrung hatten. Ausserdem wurden Apparate zweier Verwicklungsstufen verwendet : (*a*) ein einfacher Radioempfänger, und (*b*) ein kompliziertes Radarsignalesimulierungsgerät. Schliesslich wurden zwei Hauptmassstäbe bezüglich der Ausführung der Fehlereingrenzung gebraucht : (*a*) die benötigte Zeit, um festzustellen, in welchem Bestandteil ein Fehler zu finden war, und (*b*) das Prozent der Gesamtzahl möglicher Fehler, das in jeder Dreiminutenperiode eliminiert wurde.

Wie von theoretischen Hypothesen vorausgesehen war, war eine der neuen Methoden, die rationale Impulsstrom-Methode, in vielen Fällen der bisherig Normalmethode beträchtlich überlegen. Eine bedeutende Verminderung (um 40 v. H.) in Fehlerlokalisierungszeit wurde z. B. erreicht, wenn erfahrene Männer auf den komplizierten Apparat (das Radarsignalesimulierungsgerät) arbeiteten.

Dazu wurden durch Gebrauch dieser neuen Methode in den folgenden Fällen bedeutsame Erhöhungen im Prozent entfernten Fehlmöglichkeiten gewonnen : (*a*) wenn unerfahrene Männer Ausbesserungsversuche auf einen einfachen Apparat machten, (*b*) wenn erfahrene Männer verwickelte Schwierigkeiten in einem einfachen Apparat lokalisierten, und (*c*) wenn erfahrene Männer Fehlergrenzung auf einen komplizierten Apparat ausführten.

Man fand die anderen neuen Methoden—die Betandteilgruppe-Methode und die Leitungengruppe-Methode—der Normalmethode gegenüber auch vorzüglich. Ihre Ueberlegenheit war aber nicht im selben Grad auffallend, wie die der rationalen Impulsstrom-Methode. Die durch die neuen Methoden erreichbaren Vorteile scheinen leicht die Probleme zu überwiegen, die im Durchführen eines neuen, darauf gegründeten Zusammensetzungsplanes zu begegnen wären.

References

Edwards, A. L., 1954, *Statistical Methods for the Behavioral Sciences* (New York : Rinehart).

Folley, J. D., and Altman, J. W., 1958, Factors Influencing the Maintenance of Electronic Equipment (The U.S. Naval Training Device Center, Port Washington, New York).

Jonckheere, A. R., 1954, A Test of Significance for the Relation Between m Rankings and k Ranked Categories. *British Journal of Statistical Psychology* 7, 93–100.

Lindquist, E. F., 1953, *Design and Analysis of Experiments in Psychology and Education* (New York : Houghton-Mifflin Co.).

McKendry, J. M., Corso, J. F., and Grant, G., 1959, An Experimental Investigation of Equipment Packaging for Ease of Maintenance (Port Washington, New York : *The U.S. Naval Training Device Center Technical Report No.* 330–2).

McKendry, J. M., Grant, G., Corso, J. F., and Brubaker, R., 1959, Maintainability Handbook for Electronic Equipment Design (Port Washington, New York : *The U.S. Naval Training Device Center, Technical Report No.* 330–5).

Miller, R. B., Folley, J. D., and Smith, R., 1953, Systematic Troubleshooting and the Half-Split Technique (Chanute AFB, Illinois : *Human Resources Research Center Technical Report No.* 53–21).

Miller, R. B. and Stebodnick, E. B., Research for Experimental Investigation of Transferable Skills in Electronic Maintenance AFPTRC *Technical Report* 58–2, ASTIA Document No. AD 152–104. January 1958.

Ray, W. S., 1960, *An Introduction to Experimental Design* (New York : The Macmillan Co.).

Rulon, P. J., Schweiker, R. F., and Gilbert, M. C., 1958, Electronic Maintenance Training : Experimental Tryout of Diagnostigrams and Backtracks (Air Force Personnel and Training Research Center, Cambridge, Mass.).

Rulon, P. J., Schweiker, R. F., and Gilbert, M. C., 1958, Electronic maintenance training : experimental tryout of diagnostrigrams and backtracks. Lowery AFB, Colorado : Air Force Personnel and Training Research Center. Report dated January 1958 under Contract No. AF 18/600)–1355. (Also published by Educational Research Corp., Cambridge, Massachusetts.)

Runyon, R. P., 1957, Maintenance Problems Anticipated as a Result of Subminiaturization (New York : U.S. Naval Training Device Center Report No. 20–05–23–2).

Siegel, S., 1956, *Nonparametric Statistics for the Behavioral Sciences* (New York : McGraw-Hill).

Standler, L. S., Popham, W. J., and Fattu, N. A., 1956. A Review of Troubleshooting. ASTIA Report No. 134903. (Office of Technical Services, U.S. Dept. of Commerce, Washington, D.C.)

ERGONOMICS RESEARCH SOCIETY

ANNUAL CONFERENCE 1960

ABSTRACTS OF PAPERS

INERT GAS NARCOSIS

By P. B. Bennett

Royal Naval Physiological Laboratory, Alverstoke

When mixtures of an inert gas and oxygen are breathed at pressure a narcosis characterized by euphoria, retardment of the mental processes and impairment of neuromuscular coordination, is produced. Some connection has been found between this narcosis and brain activity. Electroencephalograph recordings have shown an abolition of the alpha blocking response. A correlation between a maintained change of the fusion frequency of flicker and abolition of alpha blocking has also been found when breathing oxy-nitrogen, oxy-argon and oxy-helium mixtures. The time to the onset of these changes has been shown to be inversely proportional to the square of the pressure. It may be possible to use these techniques for the selection of subjects susceptible not only to narcosis but also decompression sickness.

Studies are continuing to verify a new formula enabling deduction of the time to the threshold for any oxygen/inert gas mixture, once the constant "K" has been determined, from $P\sqrt{T}=K$, for one gas. This gives further support to the rôle of oil solubility and molecular weight in inert gas narcosis and suggests an adsorption process for the production of the narcosis. 'Additive' and 'Acclimatization' effects of frequent compressions are also being investigated.

SOME ERGONOMIC ASPECTS OF THE RESPIRATOR

By M. Ainsworth

Chemical Defence Research Establishment, Porton

A respirator limits the field of vision, restricts acoustic communication, increases the resistance to breathing, and imposes a heat load. Therefore wearing a respirator continuously means a constant reduction in the combat efficiency of the soldier. The alternative procedure of donning the respirator only at the start of a gas attack may result in casualties because of delays in warning systems and in the actual adjustment of the mask. There is therefore a choice of risking an eventual loss in overall fighting power due to such casualties, or of accepting the restriction in work rate and fighting capacity of troops imposed by continuously wearing the respirator. Some of the factors likely to influence this choice are considered and their significance evaluated.

LABOUR EXPENDITURE IN BUILDING CONSTRUCTION

By W. J. Reiners
D.S.I.R. Building Research Station

In the course of research at the Building Research Station to improve the efficiency and economy of building many studies have been made of the man-hours required for site construction, ranging from national surveys of productivity in house building to detailed analysis of the components of particular building to detailed analysis of the components of particular building operations. The paper draws on these sources to give a general account of the labour requirements of building operations and the causes for its variation.

Variations on labour expenditure are shown to arise mainly from differences in the organization and supervision of building firms rather than from local site conditions or the quality of labour. The causes of variations between firms are examined and measures given of the influence of various organizational and other factors.

Repetitive work in building shows the typical ' improvement ' pattern ; the characteristic pattern of improvement is illustrated for housing sites. Examples are given of the influence of gang size, organization, changes in incentive payments, etc. and methods of observation are discussed.

Individual productivity in craft operations is shown to depend more on motion patterns than speed of movement, and the character of individual variations in productivity is described.

THE EFFECTS OF VIBRATION ON VISION

By D. H. Drazin
R.A.F. Institute of Aviation Medicine, Farnborough

Recent research into the effects of vibration on vision is reviewed with special reference to military applications. The goal of this research has been to elucidate the relation between the loss of visual acuity suffered by the crew of high-performance vehicles and aircraft and the physical characteristics of the vibration to which they are exposed. While the vibrations of vehicles and aircraft are generally complex in form involving rotational as well as linear components, it is suggested that the study of sine-wave vibrations currently offers the most fruitful approach to the above goal.

Four factors contributing to the loss of acuity are discussed and illustrated by original experiments : (1) movement of the image over the retina, (2) the rôle of eye movements in tending to maintain or disrupt fixation, (3) the transmission of vibration through the body to the head and eyes, and (4) motivational factors associated with vibration stress.

THE ACCURACY OF SETTING OF MACHINE TOOLS BY MEANS OF HANDWHEELS AND DIALS

By E. N. Corlett
University of Birmingham

For the engineer, the thing he requires to know is the limits of accuracy of the processes he specifies. Most of the mechanical side of machine behaviour is

now predictable to close limits, but the operator is not. When setting a machine the accuracy achieved is important, the time required to set it relatively unimportant. Work attempting to assist machine designers in this respect should thus use accuracy as a criteriion.

In some work involving the study of handwheels a 2^7 half-replicate experiment, using two subjects, was first performed. From an analysis of this, randomized-block experiments were set up to define the surfaces for some of the significant interacting factors from the first experiment. Four such surfaces were defined. Two of them related dial diameter and line thickness, on ordinates showing the accuracy achieved and overall setting time. Two more surfaces related handwheel diameter and spindle resisting torque, again with the same two ordinates.

ON THE LAWS GOVERNING THE TIMES NEEDED FOR PIN-MOUNTING, AS A FUNCTION DIAMETER AND TOLERANCE

By J. F. Schouten, J. Vredenbregt and J. J. Andriessen

Instituut voor Perceptie Onderzoek
Insulindelaan 2, Eindhoven, The Netherlands

Extensive experiments were carried out on a set of calibre target holes of 1, 2, 4, 8 and 16 mm and calibre pins, fitting with tolerances 50 per cent, 25 per cent, . . . etc. down to a minimum absolute tolerance of 0·004 mm.

The laws found to govern the relation between mounting times and the diameters of pins and holes also permitted one to interpolate and extrapolate the rather erratically-jumping values in the Work-Factor Assembly Table for Closed Targets.

Suggestions are given regarding a more appropriate choice of classes of pin-and-hole diameters in such practical assembly tables.

The parameters in the formula depend in a decidedly different way on the rate at which mounting is performed. The normally used multiplicative factor, to correct for differences in rate, is not applicable in this case.

THE *DONDERS*, A 20-CHANNEL REACTION TIME AND REACTION QUALITY METER

By J. F. Schouten†, B. Lopes Cardozo†, J. Domburg‡ and F. F. Leopold†

† Instituut voor Perceptie Onderzoek, Insulindelaan 2, Eindhoven, The Netherlands

‡ Philips Research Laboratories, Eindhoven, The Netherlands

(This paper was read by Mr. F. F. Leopold)

The *Donders* measures and records the times and qualities of the reactions of 20 subjects to a common stimulus. It is equipped for distinguishing 29 different stimuli and reactions.

After each trial the data obtained are recorded on either punched tape or on an electrical typewriter. The minimum time between two trials is three seconds.

The tape is used for further automatic processing.

Particular attention has been given to adequate ergonomic design of the operator's desk. This desk embodies three panels : the presetting panel, the monitoring panel, and the checking panel.

DISCOMFORT LIMITS AS CRITERIA IN FITTING TRIALS

By J. Chris Jones

Industrial Design Ergonomics Laboratory, Metropolitan-Vickers, Manchester

The method of experimental trials is changed so that each dimension is varied in turn to find upper and lower discomfort limits for each dimension for each subject. Final settings are chosen to fall inside all discomfort limits. Seat dimensions obtained by this method are compared with those of Ackerblom. The limitations of the method are discussed.

JOB DESIGN FACTORS

By Louis E. Davis

Associate Professor of Industrial Engineering, University of California, Berkeley

Reported here is part of a continuing series of studies seeking to develop design criteria or decision rules for the design of jobs so as to minimize the total economic cost of production. Present studies are examining the relationship between job content and quantitative criteria of performance. The first study identified the very narrow criteria used at present in designing jobs in U.S. manufacturing industries. A second study identified experimentally unspecialized mass production jobs with a resulting threefold increase in quality and no decrease in quantity of output, plus positive changes in employee attitudes toward company and acceptance of responsibility for performance.

The latest study identifies job content and perception factors significantly correlated with quantitative criteria of effective performance in a chemical industry where skilled and unskilled jobs have been enlarged. The job factors that significantly correlated with reduced costs and improved quantity and quality of output are:

1. Full work assignment.
2. Fully specified work assignment.
3. Perception of job as important.
4. Identification of high quality needs.
5. Control over quality.
6. Identification of high performance with success in company.
7. Completion activities on part or product included in job assignment.
8. Communication with others.
9. Wide job knowledge.

In addition, when job enlargement proceeds to the point of providing a skilled job, workers seem to become positively more responsive to many problems and issues of concern to management.

THE EFFECT OF SIGNAL FREQUENCY IN A CHECKING TASK

By D. E. Broadbent
M.R.C. Applied Psychology Research Unit, Cambridge

Performance was examined on a task in which subjects had to detect any failure of a lamp to light up when a button was pressed. Thus they could test the lamp whenever they chose. Some subjects received a low incidence of signals during their half-hour period of work, others a high incidence of signals, and others again a high incidence of signals distributed over three separate lamps, each controlled by its own button.

This type of task has been suggested by other investigators as a means of measuring the degree of attention paid by a man to the stimuli he is supposed to inspect ; the rate at which he presses the button in order to have a look at the stimulus is supposed to reflect his level of attention.

It was found that the rate with which subjects worked increased steadily during their period of work, but that the number of observations of a signal necessary for it to be detected also increased. The time taken for detection also increased.

The rate of observation did not appear to be consistently affected by signal frequency, but did seem to be affected by the particular person who was administering the experiment.

Thus it does not seem that the rate with which items are inspected in a self paced situation of this type is a very good index of efficiency since the action of pressing the button may not be perfectly correlated with the observation of the signal.

THE EFFECT OF UNWANTED SIGNALS IN A VIGILANCE TASK

By W. P. Colquhoun
M.R.C. Applied Psychology Research Unit, Cambridge

Several studies of vigilance claim to have shown that the proportion of signals detected during a watch-keeping session increases in a fairly regular fashion with the absolute number of these signals occurring in a given period of time. However, in each of these studies, other 'unwanted' signals, of similar appearance to the 'wanted' ones, also appeared on the display. This means that each experimentally induced increase in the number of wanted signals resulted automatically in an increase in the probability that any signal occurring would be a wanted one. Because of this, it is not clear whether the observed effects are due to changes in the absolute number of wanted signals, or to alterations in the ratio of wanted to unwanted signals.

Since many practical monitoring tasks (e.g. inspection in industry, radar operation in the Services) do in fact involve discrimination between two kinds of signal rather than absolute detection of one kind, it is important to discover the exact rôle played by the (usually many) unwanted signals in the detection of the (usually few) wanted ones. A detection task was studied in which the probability of any signal being a wanted one could be varied independently of the absolute number of wanted signals presented. The results suggest that wanted signal probability is of greater importance in determining detection efficiency than the actual rate at which these signals occur ; and that the effects of alterations in this probability are maximal at the periphery of the display area being searched.

INDIVIDUAL DIFFERENCES IN THE EFFECT OF MOTIVATION ON AN INSPECTION TASK

By D. W. J. Corcoran
M.R.C. Applied Psychology Research Unit, Cambridge

Two groups of subjects were tested under conditions of low and high motivation on an inspection task. The task required the subject to cross out as many letter e's as he could in fifteen minutes, in a passage of prose. The subjects had previously taken a personality questionnaire of introversion-extraversion and neuroticism.

It was found that under these conditions speed was greater in the high motivation group, but accuracy was somewhat poorer.

The missed signals tended to fall into three kinds, viz. the e's in THE and e's which were associated with no vowel sound accounted for about 95 per cent of the omissions, whereas e's which were associated with a vowel sound only accounted for about 5 per cent of omissions.

The better scores in speed in the low motivation group were obtained by the introverts, who also tended to be less accurate. In the high motivation group speed tended to be negatively correlated with introversion. Inspection of the data revealed that in the high motivation group, extraverts only were superior to low motivation performance, whilst introverts apparently derived no benefit from increased motivation.

Several implications of these results are discussed in reference to selection for work on inspection tasks, and introduction of incentives.

PACED VERSUS UNPACED VIGILANCE

By R. T. Wilkinson
M.R.C. Applied Psychology Research Unit, Cambridge

Paced and unpaced forms of the same vigilance test have been compared. Eight hundred presentations of the display were made over an hour's watch: thirty two of these, occurring randomly, contained a critical feature or signal which had to be noticed and reported. In the paced version the successive displays were presented automatically every four and a half seconds ; in the

unpaced the subject evoked each successive display by pressing a key in his own time ; a limitation here was that rates of presentation consistently higher or lower than a four and a half second average were prevented.

In two separate experiments no difference was found between the two versions in terms either of total signals missed during the hour's test or of the decline in signal detection during this period.

Discussion centres on the implications of this result for theories of vigilance and especially the application of Broadbent's filter theory.

PERCEPTION AND ALERTNESS

By E. Elliott

Admiralty Research Laboratory, Teddington

Many of the published results of laboratory studies of vigilance do not indicate what occurs in practical watchkeeping tasks. Thus, for example, the 'fatigue' effects which Mackworth studied intensively, are seldom found in the more important military watchkeeping tasks. Again the relationship between rate of stimulation and probability of response appears to be more nebulous in protracted practical tasks than it is in a more closely controlled laboratory study such as Deese's. Furthermore we find that performance in practical tasks is generally very poor compared with what seems to be theoretically possible. This effect can be seen also in some of the well-known laboratory results, but there has been little discussion devoted to it.

This paper reformulates current ideas of perceptual organization in watchkeeping, and attempts to explain why there are discrepancies of the kind just mentioned. Particular emphasis is placed upon the type of investigation which needs to be pursued in order to resolve our difficulties.

THE BEST TIME OF DAY FOR OPERATIONS

By W. S. S. Ladell

Chemical Defence Experimental Establishment, Porton

Under certain circumstances a Commander can choose the time of day he will mount an operation. In addition to the tactical considerations there are physical and physiological factors to be taken into account, but hard and fast rules cannot be laid down. Obvious factors, operative in all climates, are the light, the direction of the sun or moon, the probable noise level, the chance of a morning mist or heat haze, and the physiological and psychological state of the men ; in a temperate climate these can probably be assessed from experience. In a tropical or desert climate the problem is more complex ; physical and physiological factors interact, diurnal changes in the temperature and humidity of the environment react on the man, the radiation load is important, also the likelihood at a given time of day of wind or rain. These factors affect his body temperature and so indirectly his maximum work capacity ; the nature of the task therefore becomes important ; different considerations will apply for a short sharp assault or a long forced march. Sweat rate must be considered, especially if water economy is important : the

possibility of sweat gland fatigue must be kept in mind ; the best conditions for rapid physical recovery after effort may be required in a phased action. In a static situation the timing of sleep and food must be considered ; early to rise if not accompanied by early to bed will cause a deterioration in fitness. Certain of these factors are considered in detail and examples of their action are given from observations made in the field on the serving soldier, with supporting observation from laboratory experiments.

TRENDS IN METHODOLOGY IN MILITARY ERGONOMICS

By E. T. Renbourn

Clothing and Stores Engineering Establishment, Farnborough, Hants.

Three investigations in military ergonomics are briefly described:

(*a*) *Assessment of Functional Clothing.* Suitable functional clothing is necessary for the soldier to carry out his varying duties. Trials are designed to compare the new and the old under both laboratory and operational conditions on fully accoutred men. Statistical design takes the form of random blocks, with variance analysis of objective data. Objective methods include measurement of sweat loss, change of body temperature and pulse rate, etc., together with water uptake by the garment layers. A simple subjective sensation score is made during replicate experiments and data analysed by appropriate methods. A preference questionnaire is taken at the end of a trial. Statistical significance must finally be converted into military significance.

(*b*) '*Fitness*' *Test.* A preliminary investigation was made into the hypothesis that soldiers working together on trials for several weeks may be as good assessors of military ' fitness ' as are objective physical fitness tests. Results showed a significant group subjective ranking of ' fitness.' This differed from a significant ranking of ' popularity '. A relationship was found between test and retest, and between ranking of ' fitness ' and results from two specific performance tests.

(*c*) *Emotional Stress.* Pulse rates and oral temperature were taken before (and sometimes after) activity on groups of boxers, athletes and parachutists, as well as controls. Results showed the existence of ' emotional fever ' under these conditions but no significant difference in pulse rate. In order to throw light on the mechanism, a controlled laboratory experiment is being carried out on small groups of men using a battery of tests. The methodology used is outlined.

ENGINEERING PSYCHOLOGY AT THE AEROSPACE MEDICAL LABORATORY, WRIGHT AIR DEVELOPMENT DIVISION

By Julien M. Christensen

Psychological Branch, Wright Air Development Center, Ohio

The contributions that engineering psychology is making to the overall systems design and development programme at the Wright Air Development

Division will be considered. A very brief résumé of the major areas in engineering psychology will be presented.

Specific areas to be considered are:

(*a*) Design for ease of maintenance (including maintenance of vehicles and stations in space).

(*b*) Design of visual, auditory and technical displays for ease and accuracy of interpretation.

(*c*) Decision making as a function of amount, significance and nature of input.

(*d*) Effects of unusual environmental conditions on performance.

(*e*) The use of servo models in the description of human motor response characteristics.

(*f*) A methodology for evaluating the restrictive characteristics of proposed space suits.

(*g*) Problems of design for complex man–machine systems (with particular emphasis on the use of models in such research).

(*h*) Design of instrument panels and controls for space vehicles.

(*i*) Problems under conditions of weightlessness.

(*j*) Methods and procedures for applying engineering psychology data to actual design problems.

The above treatment will be very general and intended to give the audience an overview of the entire programme.

In addition, the details of two selected topics, namely, human performance under conditions of weightlessness and human performance as affected by long periods of confinement, will be considered.

THE EFFECTS OF AGE AND RESPONSE COMPLEXITY IN PATTERNS OF MOVEMENT

By D. G. Entwisle and Others
University of Bristol

In tasks which require movement to a target or control, in response to a stimulus, the effect of age has been observed as an increase in the time taken to initiate responses, whilst movement time itself shows no significant change; this difference increases with greater complexity of final response.

It is held that this increase in initiation time in older people is due to a limitation in the ability to overlap monitoring of movement with the process of decision making, and it is supposed that this may result in movements of older people which are less ‘ flowing rhythmic wholes ’.

Experiments have been carried out to examine differences in movement pattern in younger and older people by means of a high speed—100/flashes/second—cyclochronographic technique. Besides confirming that an increase in response initiation time does occur in older subjects when response complexity

increases, analysis of the photographic records obtained with the technique used, has shown that changes in pattern of movement do occur. Whilst overall movement time does not show a significant increase in older subjects, changes in acceleration pattern occur, which seem to relate to the theory mentioned as well as to the physiological work on fine muscular control.

THE ASSESSMENT OF PROCESS-OPERATOR SKILLS

By E. R. F. W. Crossman and John Spencer
Reading University

The process operator, a relatively unskilled man who operates complex continuous-flow production plant, is of increasing importance in modern industry. As yet his work and skill have received little attention from research workers, and there are no established methods for assessing individual ability or contribution to plant performance. The paper outlines some interim results of a research project on this topic.

A study of settings made on a simple automatically controlled plant showed that individual operators used consistently different patterns of behaviour, and that the differences were more marked under difficult operating conditions. As no definite criterion of ' good ' operation had been laid down by the management the relative ability of operators could not properly speaking be measured. Attempts are being made to correlate the observed differences with individual characteristics.

More detailed study of operator behaviour on a smaller manually-controlled automatic machine showed that better operators as assessed from production records had a more regular pattern of monitoring and adjustment, but there was a tendency to ' overcontrol ' giving rise to unnecessary fluctuations of quality.

A laboratory experiment on a temperature-control task gave results suggesting that number of control alterations is a better index of skill than steadiness of output ; little improvement was seen in the latter respect over several trials but the number of adjustments fell steeply. When indicator lag was introduced all subjects showed oscillation of output ; some were able to overcome this after practice, but others made little improvement.

These and other observations suggest that control behaviour should be divided into at least four sub-headings:

1. Regulation or stabilization of plant running.
   - (*a*) Maintaining set conditions in the face of disturbance.
   - (*b*) Changing to new settings efficiently.
2. Optimizing plant performance.
   - (*a*) Finding the best settings for efficient working.
   - (*b*) Minimizing time lost and cost of breakdowns.

Implications of the work for personnel selection and training will be discussed. It is hoped that further work may shed some light on the underlying thought and decision processes used by unsophisticated subjects in complex situations.

THE ORGANIZATION OF FIELD TRIALS

By H. S. Wolff and J. M. Adam

Division of Human Physiology, National Institute of Medical Research, London

The field trials with which the paper will be concerned are of two types. The first can be regarded as a very large-scale physiological investigation carried out on human subjects in the field. The second type concerns physiological investigations carried out on a much more modest scale but under conditions of much greater geographical isolation and experimental difficulty than the first. An example of the first type would be a metabolic survey on Army personnel, which will be described in greater detail, and of the second type, the kind of investigations which might be made on an Antarctic expedition.

The organizational problems which arise may be due to any or all of the following features which are not usually encountered by the research worker working in the laboratory:

(*a*) The use of a very large number of subjects.
(*b*) The use of subjects who have duties to perform with which the physiological investigation is likely to interfere.
(*c*) Geographical remoteness from established laboratory facilities.
(*d*) The requirement for very robust apparatus, easy to service and of inherent high reliability.
(*e*) The accumulation of very large amounts of data.
(*f*) The recruiting of sufficient and suitable staff to man the trial.

Provided that sufficient resources are available, geographical remoteness can be countered by the use of mobile laboratory facilities and by the accumulation of sufficient spare parts and stores to deal with any eventuality. The strain placed on actual instrumentation depends very much on the complexity of the measurements which are to be made. It is however possible, with modern miniaturization techniques, to manufacture apparatus of a high degree of sophistication which yet exhibit satisfactory reliability and robustness. The requirement for reliability is particularly important in the type of instrument where it is intended to establish a continuous record, over a period of some days or weeks, of a particular physiological variable. Under these conditions each measurement is unique and cannot be repeated, and therefore a reading lost due to a faulty instrument is irreplaceable. It is therefore necessary not only to carry spare instruments but also to set up a maintenance system by which each instrument is withdrawn from use at regular intervals for routine checking and a system which is capable of detecting failure very quickly. The latter requirement means that at least a proportion of the data or samples have to be processed immediately which, in the case of gas samples for instance, requires yet further laboratory facilities.

The accumulation of very large amounts of data, which is inseparable from the employment of large numbers of subjects, can be a serious embarrassment ;

it is therefore essential to design the instrumentation in such a way that the maximum amount of data reduction occurs at the point of measurement. This will be illustrated with particular reference to time and motion study. The requirements for large numbers of staff, particularly in investigations which continue throughout the 24 hours, means in general that the majority of them will be relatively unskilled. Highly qualified technical staff has therefore to be used to its best advantage, such as in the maintenance of instruments and the assessment of data.

The problems are rather different when one considers the second type of field trial where measurements may have to be made by a member of an expedition under conditions of complete isolation from the parent laboratory and minimal local facilities. Here, our experience has shown that it is extremely difficult to design instruments which can be maintained by the expedition member concerned unless the individual has played an appreciable part in their development or has technical qualifications of a high order. In addition, his subjects are likely to be other working members of an expedition with their own work to do and not necessarily amenable to the idea of being guinea-pigs. It would appear therefore that, at present, the most valuable work which can be done in these circumstances is to restrict the work to very simple observations which become valuable because they can be continued on the same subjects over a long period.

THE ENERGY EXPENDITURE OF SOME ELDERLY MEN IN HEAVY AND LIGHT INDUSTRY

By J. V. G. A. Durnin
Institute of Physiology, The University, Glasgow

As part of a large-scale survey of food intake and energy expenditure of elderly (sic) people (aged 55–75) in Scotland, measurements have been made on groups of elderly men working in heavy and light industry. The numbers are, so far, small but are being extended. Nine men (mean age 58 years) doing heavy work in the steel industry and two groups, each of 12 men (mean ages 62 and 59 years respectively), working in a motor-lorry assembly factory and in a sewing-machine manufacturing plant, have been investigated. Total energy expenditure, covering the whole 24-hour day, has been measured on each subject for seven consecutive days, by the method previously described (Durnin and Brockway 1959).

The men in the steel works were on short time (i.e. only working about four days per week) and, in any case, their routine was very episodic so that the physically severe effort was required for only perhaps one hour in the day. Nevertheless their average energy expenditure was about 3300 cal/day. The men in the so-called ‘ light industry ’ had a much more regular pattern of work and their energy output was about 2700 cal/day. These results will be discussed in relation to other studies.

Reference

Durnin, J. V. G. A. and Brockway, J. M., 1959. Determination of the total daily energy expenditure in man by direct calorimetry: assessment of the accuracy of a modern technique. *Brit. J. Nutr.*, **13**, 41–53.

REHEARSAL AND RECALL IN IMMEDIATE MEMORY

By A. F. Sanders and E. W. J. Zwann
Institute for Perception RVO-TNO, Soesterburg, The Netherlands

This paper describes some experiments concerned with the relation between length of rehearsal and performance in recall with memorized material.

The results show that increasing the rehearsal time leads to a change in some typical aspects of immediate memory.

The memory-span can be extended, the influence of retroactive inhibition is diminished, and the decay due to differences in rehearsal and recall-rate disappears.

Theoretically, we may consider that rehearsal causes a transition from the immediate storage to more permanent forms of memory. Some aspects of this transition are discussed.

HUMAN FACTOR ASPECTS IN THE DEVELOPMENT OF THE U.S.N.E.L. FLIGHT DECK COMMUNICATION SYSTEM

By John C. Webster
Human Factors Division, U.S.N. Electronics Laboratory

The U.S. Navy Electronics Laboratory was responsible for development of a two-way, portable, noise-resistant flight deck communication system. From the very beginning a team approach was used, and Human Factors people worked together with the development engineers. The Human Factors people were completely responsible for providing television, and speech and hearing techniques, and jointly responsible for systems analyses and evaluation.

Much time was spent at sea ascertaining who need say what to whom; what information did, who need to make decisions about what, etc. It was found that one of the two command centres was blind and needed closed-circuit T.V. flight deck surveillance in order to make decisions assigned to that control centre, and that voice information need pass both ways (to and from) people immersed in extremely high noise levels and who needed complete freedom of movement on the flight deck.

Commercially available T.V. was installed to view the fore, mid, and after sections of the flight deck and the forward end of the hanger deck. Noise shielding muffs were used around ears and mouth of deck personnel wearing V.H.F. transmitter receivers in a specially designed helmet. Special noise-activated automatic volume control circuitry was employed in the receiver as was 15 dB of peak clipping. Automatic-gain control of speech level was used in the transmitter. No sidetone was used (to make the users speak as loudly as possible) and limits of maximum power supplied to the ear were set (to avoid deafening effects).

The evaluation consisted of making intelligibility tests in differing operating noise levels, recording and analysing messages (for repeats, routing, information content, and for comparing to earlier methods of passing the same

information). Written questionnaires and personal tape recorded interviews were used to find weak and strong points, morale factors, and to solicit future recommendations (allow the men a grumble session).

A (15 min) coloured film report on the project will be shown and tape recordings of the messages actually passed in varying noise levels will be available (and played, time permitting).

PHYSIOLOGICAL WORK LOAD IN SOME JOBS IN THE BUILDING CONSTRUCTION INDUSTRY

By N. P. V. Lundgren

Gymnastiska Centralinstitutet, Stockholm, Sweden

In 1959, the Employers' Association of the Swedish building construction industry in connection with their 40th anniversary set up a fund for studies of physiological and medical problems of building construction work. As the first part of this programme, a survey has been made of the physiological work load of some jobs on a relatively highly mechanized building site near Stockholm.

The studies were made during the cold weather season. Eight carpenters and eight general labourers were followed up for a few days during their ordinary work with pulse rate counting (at random time intervals), and determinations of the rectal temperature and sweating rate during work. It is seen from the table below that pulse rate levels were mainly in the range of low and moderate physiological work load, according to *Christensen's* schedule for physiological work grading. However, figures for individuals were sometimes high, showing that some work operations may be quite straining for some workers even at a high degree of mechanization. The same was true for the rectal temperature. Sweating rates were, of course, low due to the cold weather and no signs of any appreciable dehydration were evident.

The studies also included continuous determinations of pulse rate (with an electrocardiographic pulse counter) and oxygen intake (with IMP) in some manual transport jobs in which concrete was transported with various types of wheelbarrows. In these cases, the physiological work load was found to reach high levels. Special studies are therefore at present taken up in order to find out if there are any possibilities to design less straining working methods and more efficient equipment for these work operations.

Results will also be presented from interviews with the workers carried out in connection with the physiological work studies.

Observations during the work day	Carpenters	General labourers
Pulse rate (random sampling):		
Daily averages, mean	95·1	94·9
Daily averages, range	81–117	75–116
Single observations, range	65–139	59–135
Rectal temperature, °C:		
Single observations, range	37·0–38·1	37·0–38·2
Sweating rate, ml/h:		
Daily averages, mean	121	92
Daily averages, range	33–257	56–176

ERGONOMICS RESEARCH SOCIETY

ANNUAL GENERAL MEETING

The 11th Annual General Meeting of the Ergonomics Research Society was held on 28 March 1960, at 8.30 p.m., at St. John's College, Cambridge.

There were 50 members present. Dr. T. Bedford took the chair.

After the Minutes of the last Annual General Meeting had been approved and signed, the Secretary presented his report. He said that there had been considerable discussion both in the Council and at meetings of the Society concerning future activities. There was a wide range of view regarding the relative importance of research and the communication of results of research on the one hand, and the development of the application of ergonomics in industry on the other hand. The Council had set up a small Working Party to reconsider the objects of the Society and the conditions for membership.

Opening the discussion, Dr. Floyd summarized the memorandum which had been prepared for the Council's consideration. The Working Party considered, and the Council agreed, that the objects of the Society should not be changed, but it was felt that there should be some re-interpretation of the conditions required for membership. A proposal favoured by the Working Party was to create a Fellowship of the Society and suitable members could be elected as Fellows. Conditions for admission as Ordinary Members could continue unchanged, as stated at present in the Rules, but these conditions could legitimately be interpreted a good deal more liberally. There was a lively discussion of these suggestions, and several speakers expressed doubts as to the wisdom of creating an extra class of membership. Professor Müller, however, was apprehensive lest the widening of the membership would mean a lowering of the standard of the meetings which might tend to become mainly educational in character. Varied views were expressed as to the rôle of the Society, with some members emphasizing the importance of research and others dissenting.

The Chairman, in closing the discussion, pointed out that the proposals by Dr. Floyd had not as yet been agreed to by the Council, but that the views expressed at the meeting would be of the greatest value to the Council in guiding them to make further recommendations.

The Treasurer then presented his report, which was circulated. Mr. Murrell said, sound as the finances of the Society were, it was unlikely that substantial balances would be available at the end of the year. The reason was that although the subscription had been raised, the amount which would be available for running the Society would now be much less than in the past, the difference being due to the inclusion of the Journal with the subscription. The Council considered, however, that the reserves of the Society were sufficiently good to justify this course of action.

The membership of the Society was described by Dr. Hellon. There were at present 211 Ordinary Members and 24 Affiliated Members. He put forward a list of candidates for membership, all the names being approved by the meeting.

Mr. Welford reported on the Journal *Ergonomics*, describing the difficulties caused by the printing strike which had effects lasting many months. It was hoped to be able to catch up on arrears of work in the course of this year, enabling Volume 3 to be completed.

The Officers and Members of the Council were then elected for the year 1960–1961. These were as follows :

OFFICERS

Hon. Gen. Secretary
O. G. Edholm

Hon. Gen. Treasurer
K. F. H. Murrell

Hon. Membership Secretary
R. F. Hellon

Hon. Conference Secretary
S. Griew

Chairman of Council
T. Bedford

Members of Council
J. E. Cotes,
E. R. F. W. Crossman,
L. V. Green,
A. H. Jones
Miss A. D. K. Peters,
W. T. Singleton,
Miss I. M. Slade,
D. Wallis.

The Secretary reminded members that at the last Annual General Meeting there had been strong support that there should be an Overseas Member of the Council. He was glad to report that Professor E. H. Christensen, of the Fysiologiska Institutionen, Stockholm had agreed to allow his name to be put forward. Professor Christensen had been proposed by Dr. Floyd, and seconded by Mr. Wallis and Miss Slade. The Chairman therefore put this proposal before the meeting ; it was unanimously approved, and Professor Christensen was elected a Member of the Council.

ERGONOMICS RESEARCH SOCIETY ANNUAL CONFERENCE 1961

This will be held in Bristol from April 17th-20th

Suggestions for papers or symposia on special topics will be welcomed by the Meetings Secretary (Mr. D. Wallis, Division of the Senior Psychologist, Manpower Department, Admiralty, Queen Anne's Mansions, London, S.W.1).

Anyone wishing to read a paper should submit a title as soon as possible, and in any case not later than October 31st, 1960.

INSTRUCTIONS TO CONTRIBUTORS

1. Articles for publication should be sent to the General Editor or to any Member of the Editorial Board.

2. Papers must be in English, French or German. Every paper must be accompanied by a brief summary, and contributors are asked if possible to supply summaries in all three languages.

3. Authors should submit a typescript, double-spaced on one side of the paper only. Footnotes should be avoided. Summaries, tables and legends for diagrams should be typed on separate sheets. Authors must ensure that the lay-out of mathematical and other formulae is clear. The typescript must represent the final form in which the author wishes the article to appear. The cost of any alteration in proof other than printers' errors may be charged to the author.

4. Diagrams should be drawn in black ink on white card or tracing paper. They should normally be sufficiently large to allow reduction in printing and the lines should therefore be bold. **All lettering should be up to draughtsmanship standard, suitably drawn in Indian ink to allow for reduction in size.** No charges are made for reproducing tables, diagrams or half-tone illustrations, but diagrams not suitable for reproduction without redrawing may be redrawn at the Author's expense.

5. References in the text should be indicated by author's name followed by the date. They should be listed alphabetically at the end of the paper in the style illustrated by the following examples:

BARTLETT, F. C., 1943, Fatigue following highly skilled work. *Proc. roy. Soc.* B, 131, 247–254.

BEDFORD, T., 1948, *Basic Principles of Ventilation and Heating* (London: H. K. Lewis).

LE GROS CLARK, W. E., 1954, The anatomy of work. In *Symposium on Human Factors in Equipment Design* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis). Pp. 5–15.

Abbreviations should be as in the *World List of Scientific Periodicals.*

6. **Consideration for publication will gladly be given to papers which have previously had a limited circulation as research reports.** Submission of a paper implies, however, that it has not been published and will not be published elsewhere without the permission of the General Editor and the Publishers. Copyright in material accepted for publication is retained by the Journal, and reproduction in whole or in part is forbidden except under the terms of the *Fair Copying Declaration* of the Royal Society or with the written permission of the Publishers.

7. Authors will receive 25 copies of their contributions without charge. Additional copies may be ordered at the time of returning proofs. Prices for additional copies may be obtained from the Publishers.

ERGONOMICS

Volume 3 July 1960 Number 3

Contents